

Residual Prophet Inequalities

Jose Correa¹, Sebastian Perez-Salazar^{2,3}, Dana Pizarro⁴,
Bruno Ziliotto^{5,6}

¹*Department of Industrial Engineering, Universidad de Chile, Chile.

²Department of Computational Applied Mathematics and Operations
Research, Rice University, USA.

³Ken Kennedy Institute, Rice University, USA.

⁴Department of Information, Operations and Decision Sciences, TBS
Business School, France.

⁵Toulouse School of Economics, Université Toulouse Capitole, France.

⁶Institut de Mathématiques de Toulouse, CNRS UMR 5219, France.

Abstract

A typical goal for a gambler facing an online selection problem is to choose the best element, particularly in comparison to the best hindsight-optimal selection. This objective is nontrivial when there is high competition for top elements. For example, highly skilled job candidates are often recruited by top companies, leaving less competitive companies with the remaining candidates. This phenomenon introduces nontrivial correlations among the remaining candidates; hence, a gambler facing this problem with misaligned beliefs might act sub optimally if she expects a top element to still be available for selection. Motivated by these considerations, we introduce the *residual prophet inequality* (**k**-RPI) problem. In the **k**-RPI problem, we consider a finite sequence of **n** nonnegative independent random values with known distributions and a known integer $0 \leq k \leq n - 1$. Before the gambler observes the sequence, the top **k** values are removed from the sequence whereas the remaining **n** - **k** values are streamed sequentially to the gambler. Upon observing a value, the gambler must decide irrevocably if to accept/reject a value without the possibility of revisiting past values. We study two variants of **k**-RPI, according to whether the gambler learns online of the identity of the variable that he sees (FI model) or not (NI model). Our main result is a randomized algorithm in the FI model with *competitive ratio* of at least $1/(k + 2)$, which we show is tight. Our algorithm is data-driven and requires access only to the **k** + 1 largest values of a single sample from the **n** input distributions. In the NI model, we provide a similar algorithm that guarantees a competitive ratio of $1/(2k + 2)$. We further analyze independent and identically distributed instances when **k** = 1. We build a single-threshold algorithm with a

competitive ratio of at least 0.4901, and show that no single-threshold strategy can get a competitive ratio greater than 0.5464.

Keywords: Prophet inequalities, Competitive ratio, Online algorithms

1 Introduction

The prophet inequality is a classical model in optimal stopping theory [1–3]. In its simplest form, a finite sequence of n independent and nonnegative random variables X_1, \dots, X_n is observed sequentially by a gambler. Upon observing the i -th value X_i , the gambler has to irrevocably accept the value and stop the process or reject the value and observe the next value in the sequence, if any. The gambler’s goal is to devise an online algorithm that maximizes the expected accepted value. The quality of an algorithm is measured by means of the *competitive ratio* which is the fraction between the expected value obtained by the algorithm and the expected optimal offline value $\mathbb{E}(\max_i X_i)$, the so-called prophet value. The competitive ratio measures the price paid by the gambler who does not know the entire sequence of values upfront. Surprisingly, [3], showed that a simple single-threshold rule guarantees a competitive ratio of at least $1/2$ and this is tight. Prophet inequalities have received renewed attention due to their applicability in posted price mechanisms and auction theory [4–6] and have become a cornerstone modeling tool for online algorithms in Bayesian scenarios and resource allocation [7–9].

In this work we introduce the *residual prophet inequality* (k -RPI) problem: For a fixed integer $0 \leq k \leq n - 1$, the k variables corresponding to the top k realizations in the sequence X_1, \dots, X_n are removed before the gambler observes the sequence. The gambler’s goal is to maximize the expected accepted value among the remaining $n - k$ variables. Since some variables have been removed and the gambler will only observe the remaining ones, two different models can be considered, depending on whether the gambler knows the identity of the observed variable at each time or not:

Full-information (FI) In this version, the gambler observes the $n - k$ variables sequentially and upon observing a value, he also observes the identity of the variable.

No-information (NI) In this version, the gambler only observes the $n - k$ remaining values after removing the largest k values.

Algorithms designed for NI k -RPI imply algorithms for FI k -RPI. Regardless of the information model (FI or NI), the gambler can only hope to accept a value comparable to the expectation of the largest value of the $n - k$ non-removed values. This is the same as the expectation of the $(k + 1)$ -th largest value in the original sequence of n values, that is, the expectation of the $(k + 1)$ order statistics $\mathbb{E}(X_{(k+1)})$.¹

Given an information model, the competitive ratio of an algorithm for k -RPI is the ratio between the expected value of the algorithm and $\mathbb{E}(X_{(k+1)})$. Hence, a competitive

¹We assume that the order statistic of the variables X_1, \dots, X_n are ordered as $X_{(1)} \geq \dots \geq X_{(n)}$. Note that this ordering is the reverse of the convention commonly used in the literature.

ratio of γ for NI k -RPI implies a competitive ratio of γ for FI k -RPI. Likewise, hard instances for FI k -RPI imply hard instances for the less informative model NI.

The k -RPI problem can be regarded as a *robust version* of the classical prophet inequality problem (case $k = 0$), where high values are impossible to obtain due to exogenous factors. The k -RPI problem is very general and naturally relates to problems such as the postdoc problem [10, 11], which have applications in hiring problems [12–14]. Specifically, one could imagine a gambler attempting to hire an employee in a highly competitive market where the top candidates are hired by leading companies, leaving the gambler to select the best applicant from the remaining pool (see also, [15]). Another related application is in advertising. Several platforms (e.g., YouTube, Spotify, Canva, Pandora) offer both a free version supported by ads and a premium version without ads. Essentially, users paying the premium opt out from observing ads, leaving the platform to focus on advertising the high-value users among the remaining free users.

In contrast to the classic prophet inequality problem, the observed values in k -RPI are correlated. The following example demonstrates that correlation plays a significant role in k -RPI, rendering the single-threshold solutions from [3, 16] unsuitable for direct application to k -RPI.

Example 1. Consider the following instance of k -RPI with $k = 1$, $n = 3$, $X_1 = 1$ with probability (w.p.) 1 and X_2 and X_3 both independent and identically distributed (i.i.d.) taking value $1/\varepsilon^2$ w.p. $\varepsilon < 1/2$, and 0 otherwise.

The quantity $\mathbb{E}(X_{(2)})$ is given by

$$\mathbb{E}(X_{(2)}) = \varepsilon^2 \varepsilon^{-2} + 2\varepsilon(1 - \varepsilon) = 1 + 2\varepsilon - 2\varepsilon^2.$$

Let us analyze the performance of the strategy with the single-threshold solutions from [3, 16], that is $\mathbb{E}(X_{(2)})/2$ and the median of $X_{(2)}$. For the former, a case analysis shows that the gambler gets 0 when $X_2 = X_3 = 0$, which happens with probability $(1 - \varepsilon)^2$ and gets 1 with the remaining probability. Thus, the gambler gets in expectation:

$$\mathbb{E}(\text{Alg}) = 1 - (1 - \varepsilon)^2 = 2\varepsilon - \varepsilon^2,$$

and therefore, we have

$$\frac{\mathbb{E}(\text{Alg})}{\mathbb{E}(X_{(2)})} = \frac{2\varepsilon - \varepsilon^2}{1 + 2\varepsilon - 2\varepsilon^2} < \frac{1}{2}.$$

Moreover, $\frac{\mathbb{E}(\text{Alg})}{\mathbb{E}(X_{(2)})} = \mathcal{O}(\varepsilon)$ and then the gambler cannot guarantee a constant factor of $\mathbb{E}(X_{(2)})$.

On the other hand, we see that the median of $X_{(2)}$ is 0, so using any threshold between the median and $\mathbb{E}(X_{(2)})$ will not produce a different result.

In fact, any strategy that accepts value 1 is ineffective because, once the value 1 has been observed, the expectation of the second variable is $1/((2 - \varepsilon)\varepsilon) \gg 1$. Such a positive correlation between the two observed variables is what makes classic strategies fail.

1.1 Results and technical contributions

The previous example illustrates that correlation plays a major role for k -RPI. The examples also show that traditional and well-liked thresholds such as the median or the expectation of $X_{(k+1)}$ can be arbitrarily poor choices. This is contrary to the negative correlated case where we can guarantee a competitive ratio of $1/2$ [17–19], as in the independent case. Our first contribution is a new algorithmic approach that bypasses this hardness.

Main Result [Lower bound on competitive ratio]

We show that for the full information model of k -RPI, there exists an algorithm with a competitive ratio of at least $1/(k+2)$. Our algorithm first samples one value from each input distribution and randomly selects one of the $k+1$ largest values in the sample. It then uses this value and the identities of the arrivals to perform the online selection. The randomization is independent of the input, and we note that our algorithm extends the approach of [20] for the classic prophet inequality problem. We present the details of our algorithm and its analysis in Section 3.

Our algorithmic solution for the FI k -RPI is robust in some key aspects. On one hand, it works against any arrival order making it highly applicable in online problems. On the other, by construction, it does not need to know the distributions of each variable, but only requires one sample from each variable. This is particularly important for applications in posted price mechanisms, where consumer valuation distributions are typically unknown, and only a limited amount of past sales data is available.

For the no-information model of the k -RPI, we prove that there exists a single-threshold strategy with a competitive ratio of at least $1/(2k+2)$. The idea is similar to that of the FI k -RPI model, but the algorithm uses only one of the top $k+1$ sample values—selected uniformly at random—as the threshold for making the online selection. The lack of information regarding the identities of the removed variables leads to a degradation in the competitive guarantee. Nevertheless, this guarantee can be transferred to the FI k -RPI model, showing that it is possible to achieve a constant-factor approximation of $\mathbb{E}(X_{(k+1)})$ using a single-threshold strategy.

Next, we prove that our main result for the FI model is tight.

Tightness [Upper bound on competitive ratio]

For any information model of k -RPI, there is no algorithm that has a competitive ratio larger than $1/(k+2)$. To provide this negative result, we construct a hard instance in the FI k -RPI model; which will imply the negative result for the model with less information. Our instance extends the hard instance for the classic prophet inequality problem: it consists of a sequence $X_1, \dots, X_{2(k+1)}$ of two-point distributions, where each $X_i \in \{0, a_i\}$. The values a_i are positive and increase rapidly with i , while the events $X_i = a_i$ occur rarely for i larger than $k+1$. We provide the details in Section 4.

Our hard instance unveils that part of the hardness of the k -RPI problem stems from the *order* in which the values are observed by the gambler. For every information model of k -RPI, if the gambler observes the values in random order (RO), then there is an algorithm with a competitive ratio at least $1/e$. The algorithm is a straightforward application of the secretary problem algorithm [21–24] and we defer

the details to Appendix A. Our last result is an exploration to the independent and identically distributed (i.i.d.) k -RPI problem where X_1, \dots, X_n are drawn from the same distribution.

Additional result [i.i.d. k -RPI, $k = 1$]

The previous observation implies that for i.i.d. random variables and arbitrary k , we can always guarantee a factor $1/e$ using the classic secretary algorithm. This shows a stark difference between k -RPI and its i.i.d. counterpart and indeed even for small k this improves upon our tight factor of $1/(k+2)$ for k -RPI. Therefore, it is interesting to explore the gap between the i.i.d. and the independent versions of the problem, even in the case $k = 1$. We prove that for both information models of 1-RPI, if the values X_1, \dots, X_n are i.i.d., then there exists an algorithm with a competitive ratio 0.4901. Our algorithm here is more standard. We propose a single-threshold strategy for NI 1-RPI which determines the threshold τ via $\Pr(X \geq \tau) = q$, where q is an input quantile. Our analysis follows a quantile-based approach, expressing both the expected value of the algorithm and the optimal value $\mathbb{E}(X_{(2)})$ as functions of quantiles. By comparing their ratio, we derive a lower bound that depends solely on q . Optimizing over q yields the desired result. We also establish that no single-threshold strategy can achieve a competitive ratio greater than 0.5464 in any information model. This result shows that the optimal competitive ratio of $1 - 1/e$ for single-threshold strategies [2, 25], attained when $k = 0$, cannot be recovered for $k \geq 1$. We present the details in Section 5.

1.2 Related Literature

The prophet inequality problem, as introduced by Krengel and Sucheston [1], was resolved by using a dynamic program that gave a tight approximation ratio of $1/2$. [3] later proved that a single-threshold strategy yields the same guarantee; this also showed that the order in which the variables are observed is immaterial. The renewed interest in prophet inequalities is due to their relevance to auctions, specifically posted priced mechanisms (PPMs) in online sales [26–30]. It was implicitly shown by [27] and [29] that every prophet-type inequality implies a corresponding approximation guarantee in a PPM, and the converse is true as well [5].

The closest work to ours is likely that of [20], where the authors used the principle of deferred decision to prove that a single sample from each distribution is sufficient to achieve a competitive ratio of $1/2$ for the classic prophet inequality. This technique has also been applied to other optimal stopping problems (see, e.g., [31, 32]).

In essence, after obtaining one sample from each distribution, [20] sets the threshold as the maximum of these samples. Although our proof for the general case is also based on this principle, the analysis is much more intricate due to the complexity of the k -RPI problem, which necessitates a more sophisticated algorithm. Specifically, for our approach to be effective, it is insufficient to simply use a threshold based on the j -th order statistic of the sample set for some fixed j . Instead, the algorithm first selects j according to a carefully chosen distribution. Moreover, in the FI model, the algorithm must discard certain elements based on their identity, even when their values exceed

the j -th order statistic. Thus, in contrast to Rubinstein's work, where j is deterministically fixed at 1, our approach introduces an additional layer of randomization, and the j -th order statistic is not exactly used as a threshold in the FI model.

There has been a growing interest in competitive versions of online selection problems [33–37]. The closest paper in this stream of literature to ours is the one by [33], where the authors consider a generalization of the prophet inequality problem with $k + 1$ gamblers. Gambler j observes the sequence after the first $j - 1$ gamblers have gone through the sequence, and they study reward guarantees under single-threshold strategies. Note that, in our case, we can imagine that there are $k + 1$ gamblers but the first k gamblers are all-mighty. These k gamblers are not strategic, hence we do not need a game-theoretic analysis, unlike in the aforementioned papers on competitive prophet inequalities.

2 Model

For $0 \leq k \leq n - 1$, an instance of k -RPI is given by a sequence X_1, \dots, X_n of nonnegative independent random variables, where X_i has cumulative density function (cdf) F_i . Nature removes k variables corresponding to the top k realizations,² and we denote by D the corresponding set of indices of the remaining variables. We consider two information models that determines what the gambler observes sequentially.

In the *full information* (FI) model, the gambler observes online the pairs $(X_i, i)_{i \in D}$. That is, the gambler observes both the value and the index of the random variable from which the value originates.

In the *no information* (NI) model, the gambler only observes online the $n - k$ values in the sequence $(X_i)_{i \in D}$. In both information models, D is unknown to the gambler upfront. Given an information model (FI or NI), the gambler wants to implement an online algorithm ALG that observes the online values according to the information model and accepts a value. Regardless of the model, and abusing notation, we denote by ALG the value accepted by the online algorithm. The expected optimal offline solution corresponds to $\mathbb{E}(\max_{i \in D} X_i) = \mathbb{E}(X_{(k+1)})$.

For $\gamma > 0$, we say that ALG has a *competitive ratio* γ if $\mathbb{E}(\text{ALG}) \geq \gamma \cdot \mathbb{E}(X_{(k+1)})$ for any input of k -RPI. For each k , we are interested in finding the largest γ_k such that there is an algorithm ALG with competitive ratio γ_k for k -RPI. Note that for $k = 0$, we have $\gamma_0 = 1/2$ [3]. We note that an algorithm with a competitive ratio γ for the NI model implies an algorithm with competitive ratio γ for the FI model.

3 Lower bound on competitive ratio

In this section, we prove our main result. We assume that the distributions F_1, \dots, F_n are independent but not necessarily identically distributed.

Theorem 1. *For the FI model, there is an algorithm for k -RPI with competitive ratio at least $1/(k + 2)$.*

²If there are several choices due to ties, Nature randomizes the choice of the k variables.

Theorem 2. *For the NI model, there is a single-threshold algorithm for k -RPI with competitive ratio at least $1/2(k+1)$.*

To prove both Theorem 1 and Theorem 2, we employ a randomized strategy. In the case of Theorem 2, the strategy is, in fact, a randomized threshold strategy. We highlight here that, as a corollary of Theorem 2, we obtain that in the FI model, there exists a threshold strategy with a competitive ratio of at least $\frac{1}{2(k+1)}$.

To understand the rationale behind the construction of our randomized strategies to prove Theorem 1 and Theorem 2, let us recall the result obtained by [20] in the classic prophet inequality setting. By drawing one sample from each distribution and taking the maximum of them as a threshold, the gambler can guarantee a competitive ratio of $1/2$. A natural adaptation of that algorithm to our setting is to consider as a threshold the $(k+1)$ -th maximum of the samples. We denote by MSA_{k+1} such a strategy.

Unfortunately, such a strategy does not guarantee any constant competitive ratio. Indeed, consider again the instance in Example 1 with $k = 1$.

The expected value of the algorithm MSA_2 is:

$$\begin{aligned}\mathbb{E}(MSA_2) &= \varepsilon^2 \mathbb{E}(MSA_2 | \tau = \varepsilon^{-2}) + (1 - \varepsilon^2) \mathbb{E}(MSA_2 | \tau \leq 1) \\ &= \varepsilon^2 \cdot \varepsilon^2 \cdot \varepsilon^{-2} + (1 - \varepsilon^2) \cdot [1 - (1 - \varepsilon)^2] \cdot 1 \\ &= 2\varepsilon - 2\varepsilon^3 + \varepsilon^4.\end{aligned}$$

Given that $\mathbb{E}(X_{(2)}) \rightarrow 1$ when $\varepsilon \rightarrow 0$, we obtain that $\mathbb{E}(MSA_2)/\mathbb{E}(X_{(2)}) \rightarrow 0$ as ε tends to zero.

To tackle this problem and prove the competitive ratios in Theorem 1 and Theorem 2, we draw one sample s_i from each distribution F_i , and consider $k+1$ algorithms. For the NI model, we analyze single-threshold algorithms, denoted by MSA_1, \dots, MSA_{k+1} , defined as follows: in algorithm MSA_i , for $i \in \{1, \dots, k+1\}$, the gambler's strategy is to set the threshold τ as the i -th largest value in the sample set, and then select the first observed value x_j exceeding τ .

For the FI model, the algorithms differ slightly for $i \in \{1, \dots, k\}$: they select the first value exceeding the threshold that does not come from the same distribution as the sample that defined the threshold. We denote these algorithms by \overline{MSA}_i for $i \in \{1, \dots, k+1\}$. It is worth highlighting that \overline{MSA}_{k+1} is identical to MSA_{k+1} .

Note that the algorithms \overline{MSA}_i , for $i \in \{1, \dots, k\}$, must determine whether the arriving value originates from the same distribution as the sample used to define the threshold, and therefore, the knowledge of the identity of each variable is necessary for the online selection.

A complete analysis of these algorithms is provided in Subsections 3.1 and 3.2, with some proofs deferred to Appendix B.

By the principle of deferred decision and following the formalism in [20], instead of considering one sample for each distribution and then looking at the real values in an online fashion, we can draw two samples from each distribution F_i , namely y_i and z_i , and then flip a fair coin to decide which is equal to s_i and which is equal to x_i . This procedure correctly generates s_1, \dots, s_n and x_1, \dots, x_n as independent draws of

F_1, \dots, F_n . From now on, we will denote by S the set of samples $\{s_1, \dots, s_n\}$ and X the set of true values $\{x_1, \dots, x_n\}$.

To analyze the performance of the algorithms, we assume that for each i , $y_i > z_i$ and we order all these samples in decreasing order, relabeling them as w_1, \dots, w_{2n} , so that $w_1 \geq w_2 \geq w_3 \geq \dots \geq w_{2n}$ ³. We say that $(w_l, w_{l'})$ is a pair, or that w_l is paired with $w_{l'}$, if they originate from the same distribution.

Moreover, for each $j \in \{1, \dots, k+1\}$ we define ξ_j as the corresponding position of the j -th value z in the sequence of w 's values. For example, if the first elements of the w sequence are given by

$$y_3 \ y_5 \ y_1 \ z_5 \ y_8 \ z_8 \ z_3 \ \dots,$$

then $\xi_1 = 4$ and $\xi_2 = 6$. Note that ξ_j can be also seen as the position at which the j -th pair (y, z) from the same distribution appears.

In the subsequent analysis, we fix specific realizations of the pairs (y_i, z_i) , which in turn determine the ξ_j and the w_i .

3.1 Proof of Theorem 1

To show Theorem 1, we consider the $k+1$ algorithms $\overline{MSA}_1, \dots, \overline{MSA}_{k+1}$ defined in Section 3, and we use them to define the randomized strategy \overline{MSA}_{RAND} as follows: (1) before the game starts, select a random number I in $\{1, \dots, k+1\}$, such that for all $i \in \{1, \dots, k\}$, $I = i$ with probability $1/(k+2)$, and $I = k+1$ with probability $2/(k+2)$. (2) Play \overline{MSA}_I .

We will prove Theorem 1 by showing that the strategy \overline{MSA}_{RAND} has a competitive ratio $\frac{1}{k+2}$. Before proceeding to the proof of Theorem 1, we need to introduce some definitions, and two technical lemmas, whose proofs have been deferred to Appendix B.

Definition 1. Let $l \in \{1, \dots, 2k+1\}$. We say that w_l is blocked if there exist $r, r' \in \{l+1, \dots, 2k+1\}$ such that $w_{r'} = y_j$ and $w_r = z_j$ for some j . We denote by m_l the smallest r that satisfies this property.

For example, if $k = 3$ and the first $2k+1 = 7$ elements of the w sequence are given by

$$y_3 \ y_5 \ y_1 \ z_5 \ y_8 \ z_8 \ z_1 \ \dots,$$

$l = 2$ is blocked, since the pairs (y_1, z_1) and (y_8, z_8) appear between the 3-rd and 7-th positions. Moreover, in this case $m_2 = 6$.

The pair (y_j, z_j) “blocks” w_l , in the sense that no matter whether $z_j = w_r$ is in X or S , no threshold below w_r can guarantee selecting the value w_l .

Definition 2. Let $l \in \{1, \dots, 2k+1\}$ and p such that w_p is paired with w_l . We say that w_l is ill-paired if $p \in \{l+1, \dots, 2k+1\}$.

³If some values are identical, Nature randomizes their order within the sequence.

That is, we say that a value w_l is ill-paired if it is paired with a value greater than or equal to w_{2k+1} . For instance, considering the same sequence as before, w_2 is ill-paired since z_5 appears before w_7 .

Definition 3. For each $l \in \{1, \dots, 2k+1\}$, we define the parameter δ_l as follows:

$$\delta_l = \begin{cases} 2^{-2k+l-1} & \text{if } w_l \text{ is not blocked and not ill-paired} \\ 2^{-2k+l} & \text{if } w_l \text{ is not blocked and ill-paired} \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 1. If the gambler plays according to \overline{MSA}_{k+1} , his expected reward is at least

$$\mathbb{E}(\overline{MSA}_{k+1}) \geq \frac{1}{2} \sum_{l=k+1}^{2k+1} \mathbb{P}(X_{(k+1)} = w_l) w_l \delta_l + \frac{1}{2} \sum_{l=2(k+1)}^{\xi_{k+1}} \mathbb{P}(X_{(k+1)} = w_l) w_l$$

Note that $\mathbb{E}(X_{(k+1)})$ is equal to $\sum_{l=k+1}^{\xi_{k+1}} w_l \mathbb{P}(X_{(k+1)} = w_l)$. Consequently, when $k = 0$, Lemma 1 recovers the result from [20] which states that \overline{MSA}_1 gives a $1/2$ competitive ratio ($\delta_1 = 1/2$ since w_1 is neither blocked nor ill-paired). The challenge when $k \geq 1$ arises from the fact that, for $k+1 \leq l \leq 2k$, the coefficient accompanying the term $\mathbb{P}(X_{(k+1)} = w_l) w_l$ may be smaller than $1/(k+2)$. In other words, for $k+1 \leq l \leq 2k$ the coefficient may be “too small”, while for $2k+1 \leq l \leq \xi_{k+1}$, it is “larger than necessary” (equal to $1/2$). This imbalance motivates the introduction of a randomization over the \overline{MSA}_i algorithms: by blending \overline{MSA}_{k+1} with \overline{MSA}_i for $i \in \{1, \dots, k\}$, we redistribute these coefficients more evenly. To analyze such a randomization, we need first a lower bound on the performance of $\overline{MSA}_i, i \in \{1, \dots, k\}$.

Lemma 2. The sum of the expected reward of the gambler playing according to \overline{MSA}_i for $i \leq k$, conditional on ξ_j for $j \in \{1, \dots, k+1\}$ is at least

$$\sum_{i=1}^k \mathbb{E}(\overline{MSA}_i) \geq \sum_{l=1}^{2k+1} \mathbb{P}(X_{(k+1)} = w_l) w_l (1 - \delta_l).$$

The coefficients accompanying the $\mathbb{P}(X_{(k+1)} = w_l) w_l$ in the above inequality are higher than those in the expression of Lemma 1 for $k+1 \leq l \leq 2k$, while they are equal to 0 for $l > 2k+1$. This supports the idea that combining algorithms enables a redistribution of coefficients. The surprising fact is that there exists a way to combine the $\overline{MSA}_i, i \in \{1, \dots, k+1\}$ in a way that all the coefficients are simultaneously higher than $1/(k+2)$, yielding the competitive factor of $1/(k+2)$. We prove this below.

Proof of Theorem 1. Let us consider the strategy for the gambler \overline{MSA}_{RAND} consisting on playing according to \overline{MSA}_i with probability $1/(k+2)$, for $i \in \{1, \dots, k\}$, and to \overline{MSA}_{k+1} with probability $2/(k+2)$.

Then, $(k+2)\mathbb{E}(\overline{MSA}_{RAND}) = \sum_{i=1}^k \mathbb{E}(\overline{MSA}_i) + 2\mathbb{E}(\overline{MSA}_{k+1})$, and by using Lemma 1 and Lemma 2, we obtain

$$\begin{aligned}
(k+2)\mathbb{E}(\overline{MSA}_{RAND}) &\geq \sum_{l=1}^{2k+1} \mathbb{P}(X_{(k+1)} = w_l) w_l (1 - \delta_l) \\
&\quad + 2 \sum_{l=k+1}^{2k+1} \mathbb{P}(X_{(k+1)} = w_l) w_l \frac{\delta_l}{2} + 2 \sum_{l=2(k+1)}^{\xi_{k+1}} \mathbb{P}(X_{(k+1)} = w_l) w_l \frac{1}{2} \\
&= \sum_{l=k+1}^{\xi_{k+1}} \mathbb{P}(X_{(k+1)} = w_l) w_l = \mathbb{E}(X_{(k+1)}),
\end{aligned}$$

where the equality holds because $\mathbb{P}(X_{(k+1)} = w_l) = 0$ for $l < k+1$.

This concludes on the proof of Theorem 1. □

3.2 Proof of Theorem 2

In order to prove Theorem 2, we use algorithms MSA_1, \dots, MSA_{k+1} defined in Section 3. Then, we define the randomized strategy MSA_{RAND} as follows: (1) before the game starts, select a number I in $\{1, \dots, k+1\}$ uniformly at random, that is, $I = i$ with probability $1/(k+1)$. (2) Play MSA_I .

Note that, unlike the randomized algorithm we used to prove Theorem 1, here MSA_{RAND} does not need access to the identity of the arriving variables. In the next proposition, we prove that MSA_{RAND} has a competitive ratio of at least $\frac{1}{2(k+1)}$.

Proposition 1. *The strategy MSA_{RAND} has a competitive ratio $\frac{1}{2(k+1)}$.*

Proposition 1 implies directly Theorem 2. Indeed, MSA_{RAND} is a randomization over single-threshold algorithms. By linearity of expectation, there exists a single-threshold strategy in the support of MSA_{RAND} that performs as well as MSA_{RAND} .

We defer the proof of Proposition 1 to Appendix B.

4 Upper bound on competitive ratio

In this section, we prove the tightness of the competitive ratio result for FI k -RPI presented in Section 3, by providing a parameterized hard distribution for FI k -RPI showing that no algorithm can have a competitive ratio of γ , with $\gamma > 1/(k+2) + \beta$ for any $\beta > 0$. This leads to a similar result for the lower-information model NI.

Theorem 3. *For each $\beta > 0$, there exists an instance with $2(k+1)$ variables such that no algorithm has a competitive ratio larger than $1/(k+2) + \beta$, regardless of the information model in k -RPI.*

Proof of Theorem 3. Let $0 < \varepsilon < 1/(k+1)$. Consider the following $2(k+1)$ random variables: $X_i = \frac{1}{\varepsilon^{i-1} \binom{k+1}{i-1}}$ for $i \in \{1, \dots, k+1\}$ and for $i \in \{k+2, \dots, 2(k+1)\}$

$$X_i = \begin{cases} 0 & \text{w.p. } 1 - \varepsilon, \\ 1/\varepsilon^{k+1} & \text{w.p. } \varepsilon. \end{cases}$$

We now prove that for the instance with these $2(k+1)$ random variables, no strategy of the gambler can attain a competitive ratio larger than $1/(k+2) + \mathcal{O}(\varepsilon)$.

Note that, as $\varepsilon < (k+1)^{-1}$, it holds $X_i < X_{i+1}$ for $1 \leq i \leq k$. That is, the deterministic variables arrive in increasing order. Indeed, for $i \in \{1, \dots, k\}$, $X_i > X_{i+1}$ if and only if $\varepsilon < \frac{i}{k-i+2}$. As $\frac{i}{k-i+2}$ is increasing in i , it is enough to have $\varepsilon < (k+1)^{-1}$.

Let us compute $\mathbb{E}(X_{(k+1)})$. To this end, note that the $(k+1)$ -th largest variable will correspond to X_i with $i \leq k+1$ if and only if exactly $i-1$ variables take the value $1/\varepsilon^{k+1}$ (because the deterministic variables arrive in increasing order with respect to their value), and it will correspond to a variable X_i with $i \geq k+2$ if and only if all variables $j \geq k+2$ take the value $1/\varepsilon^{k+1}$.

We define the random variable Y as the number of variables among $X_{k+2}, \dots, X_{2(k+1)}$ that take the value $1/\varepsilon^{k+1}$. Then, conditioning on the value of Y , we have:

$$\begin{aligned} \mathbb{E}(X_{(k+1)}) &= \sum_{j=0}^{k+1} \mathbb{E}(X_{k+1} | Y = j) \mathbb{P}(Y = j) \\ &= 1 \cdot (1 - \varepsilon)^{k+1} + \sum_{j=1}^k \mathbb{E}(X_{k+1} | Y = j) \cdot \mathbb{P}(Y = j) + \frac{1}{\varepsilon^{k+1}} \cdot \varepsilon^{k+1} \\ &= (1 - \varepsilon)^{k+1} + \sum_{j=1}^k \frac{1}{\varepsilon^j \binom{k+1}{j}} \cdot \binom{k+1}{j} \cdot \varepsilon^j \cdot (1 - \varepsilon)^{k+1-j} + 1 \\ &= 1 + \sum_{j=0}^k (1 - \varepsilon)^{k+1-j} \end{aligned}$$

Let us now compute the optimal guarantee of the gambler. First, observe that the gambler should always accept the value $1/\varepsilon^{k+1}$, as it is the highest possible one. Furthermore, since the deterministic values are strictly increasing, if the gambler sees that a deterministic value has been removed, then all remaining variables—both deterministic and non-deterministic—must have either been removed or are equal to zero. In this case, the gambler receives 0. As a result, the gambler does not gain any useful information from observing past values, allowing us to restrict to strategies of the following form: (1) stop at time i , for some $i \leq k+1$ (2) stop at the first positive value appearing after stage $k+2$.

Call ALG_i the payoff of a strategy of the form (1). Under such a strategy, the gambler will pick X_i if and only if there are at least $i-1$ variables taking a value

$1/\varepsilon^{k+1}$. That is, if Y is greater than or equal to $i - 1$. Then, under this strategy, the gambler obtains in expectation

$$\begin{aligned}\mathbb{E}(ALG_i) &= \sum_{j=i-1}^{k+1} \mathbb{E}(ALG_i | Y = j) \mathbb{P}(Y = j) \\ &= \frac{1}{\varepsilon^{i-1} \binom{k+1}{i-1}} \sum_{j=i-1}^{k+1} \binom{k+1}{j} \varepsilon^j (1-\varepsilon)^{k+1-j} \\ &= (1-\varepsilon)^{k-i+2} + \sum_{j=i}^{k+1} \frac{\binom{k+1}{j}}{\binom{k+1}{i-1}} \varepsilon^{j-i+1} (1-\varepsilon)^{k+1-j}.\end{aligned}$$

Last, consider strategy (2). This strategy gets a positive payoff if and only if all variables $X_{k+2}, \dots, X_{2(k+2)}$ are positive, which happens with probability ε^{k+1} . When this is the case, it gets payoff $\varepsilon^{-(k+1)}$. Consequently, (2) guarantees $\varepsilon^{k+1} \cdot \varepsilon^{-(k+1)} = 1$.

It follows that the optimal payoff of the gambler goes to 1 as $\varepsilon \rightarrow 0$. Moreover, we have $\mathbb{E}(X_{(k+1)}) \rightarrow k+2$ as $\varepsilon \rightarrow 0$. Consequently, for each $\beta > 0$, one can find $\varepsilon > 0$ such that no algorithm achieves a competitive ratio larger than $1/(k+2) + \beta$ in the corresponding instance. This proves the theorem. \square

5 I.I.D. Case for $k = 1$

In this section, we focus on i.i.d. instances where $F_1 = \dots = F_n$ in the case of k -RPI with $k = 1$. That is, in the sequence X_1, \dots, X_n , the maximum value has been removed. The main result of this section is the following:

Theorem 4. *For any information model, there is an algorithm for 1-RPI with a competitive ratio of at least 0.4901.*

For the rest of the section, we assume that X_i are continuous with cdf $F(\cdot)$. Furthermore, following [38], we can also assume that F is strictly increasing and infinitely differentiable. Since the gambler observes the sequence of $n-1$ values, we can assume that the maximum of the n values occurs in the last position n . Hence, the gambler faces the problem under the event $E = \{X_1, \dots, X_{n-1} < X_n\} = \{\max_{i < n} X_i < X_n\}$. Note that $\mathbb{P}(E) = 1/n$ by the continuity of F .

To prove Theorem 4, we provide a fixed-threshold strategy that computes a threshold based on quantiles $q \in [0, 1]$. That is, given $q \in [0, 1]$, the algorithm will compute $u \geq 0$ such that $q = \mathbb{P}(X \geq \tau) = 1 - F(\tau)$, and will accept the first value at least u in the observed sequence. We denote such an algorithm ALG_q . The following lemma provides a lower bound for a particular choice of quantiles q .

Lemma 3. *Let $n \geq 3$. For NI 1-RPI, if ALG_q is run with $q = \alpha/(n-1)$ for $\alpha \in [0, 2]$, then,*

$$\frac{\mathbb{E}(ALG_q)}{\mathbb{E}(X_{(2)})} \geq \min \left\{ \frac{1 - e^{-\alpha}}{\alpha}, 1 - e^{-\alpha}(1 + \alpha) \right\},$$

for any continuous cdf F .

Using this lemma, and by equating $(1 - e^{-\alpha})/\alpha = 1 - e^{-\alpha}(1 + \alpha)$, we obtain that $\alpha \approx 1.64718$ and the competitive ratio of fixed-threshold solutions is ≥ 0.4901 . We can also show that no fixed-threshold solution can obtain a competitive ratio better than 0.5463. We defer this analysis to Appendix C.3.

In the remainder of the section, we provide the details of the proof of Lemma 3.

5.1 Proof of Lemma 3

In this subsection, we provide the lower bound on the competitive ratio of ALG_q for $q = \alpha/(n-1)$. For notational convenience, we will avoid writing the subscript in ALG . The algorithm ALG computes the threshold u in advance and accept the first observed value that surpassed u . Then, the reward of ALG as a function of u is

$$\mathbb{E}(ALG) = \sum_{i=0}^{n-2} \mathbb{P}(X_1, \dots, X_i < u \mid E) \mathbb{E} [X_{i+1} \mathbf{1}_{\{X_{i+1} \geq u\}} \mid X_1, \dots, X_i < u, E]$$

By analyzing the different involved probabilities, we can find the following characterization of $\mathbb{E}(ALG)$ as a function of the quantile q :

Proposition 2. *If $q \in [0, 1]$, then*

$$\mathbb{E}(ALG) = \int_0^1 r(v) \sum_{i=0}^{n-2} \frac{n}{n-i-1} (1-q)^i \left(\min\{q, v\} - \left(\frac{1 - (1 - \min\{q, v\})^{n-i}}{n-i} \right) \right) dv,$$

here $r(v) \geq 0$ is such that $F^{-1}(1-u) = \int_u^1 r(v) dv$ which exists due to the assumptions over the cdf F .

The proof of this proposition is technical and deferred to Appendix C.1. Likewise, we can find an expression for $\mathbb{E}(X_{(2)})$ in terms of $r(v)$ from the Proposition:

$$\mathbb{E}(X_{(2)}) = \int_0^1 r(v) \mathbb{P}(\text{Binom}(n, v) \geq 2) dv.$$

Then,

$$\frac{\mathbb{E}(ALG)}{\mathbb{E}(X_{(2)})} \geq \inf_{v \in [0, 1]} \left\{ \frac{\sum_{i=0}^{n-2} \frac{n}{n-i-1} (1-q)^i \left(\min\{q, v\} - \left(\frac{1 - (1 - \min\{q, v\})^{n-i}}{n-i} \right) \right)}{\mathbb{P}(\text{Binom}(n, v) \geq 2)} \right\}$$

This last bound is instance-independent and only depends on n and q . Let $A_{n,q}(v)$ be the function in the infimum. We study the the infimum of $A_{n,q}(v)$ for the regime

$v > q$ and $v < q$ separately. The following proposition characterizes the behavior of $A_{n,q}(v)$ in both regimes for $q = \alpha/(n-1)$ and $\alpha \leq 2$. We defer the proof of this result to Appendix C.2.

Proposition 3. *For $q = \alpha/(n-1)$ and $\alpha \leq 2$, we have*

1. *If $v > q$, then, $A_{n,q}(v)$ is decreasing in v ;*
2. *If $v \leq q$, then, $A_{n,q}(v)$ is increasing in v .*

With this proposition, we obtain

$$\begin{aligned}
\frac{\mathbb{E}(ALG)}{\mathbb{E}(X_{(2)})} &\geq \min \left\{ \inf_{v \in [0,q]} \{A_{n,q}(v)\}, \inf_{v \in [q,1]} \{A_{n,q}(v)\} \right\} \\
&= \min \left\{ \lim_{v \rightarrow 0} A_{n,q}(v), A_{n,q}(1) \right\} \\
&= \min \left\{ (n-1) \frac{1 - (1-q)^{n-1}}{q}, 1 - (1-q)^{n-1}(1 + (n+1)q) \right\} \\
&\geq \min \left\{ \frac{1 - e^{-\alpha}}{\alpha}, 1 - e^{-\alpha}(1 + \alpha) \right\}
\end{aligned}$$

where in the first equality we use Proposition 3, the second equality follows by a simple calculation, and in the last inequality we use the standard inequality $1 - x \leq e^{-x}$. This finishes the proof of Lemma 3 and by setting $q = \alpha/(n-1)$. This finishes the proof of the lemma.

6 Conclusion and Final Remarks

In this work, we introduced the residual prophet inequality problem (k -RPI), a new variant of the classical prophet inequality model where the top k variables are removed before observation. Our formulation highlights the impact of correlation in sequential selection problems and demonstrates that classical single-threshold approaches are insufficient in this setting. We provided a randomized algorithm with a competitive ratio of $1/(k+2)$ for the FI model and showed the tightness of this bound. For the NI model, we give a randomized threshold algorithm with a competitive ratio of $1/(2k+2)$. Additionally, we analyzed the i.i.d. case of 1-RPI and proposed an algorithm with a competitive ratio of at least 0.4901. Furthermore, we proved that no single-threshold strategy can achieve a competitive ratio greater than 0.5464.

Since this is the first time k -RPI is introduced, our work opens up several promising directions for future research. One such direction is to investigate whether the $1/(k+2)$ competitive ratio can be achieved using a threshold-based strategy. Another natural avenue is to determine the tight competitive ratio for single-threshold strategies in the i.i.d. case of 1-RPI, and to explore how these results might extend to k -RPI for $k \geq 2$. One of the limitations of our current analysis is that it relies heavily on being able to compute probabilities under the condition of the maximum value being removed; these probabilities become intractable to handle for larger values of k . Naturally, determining the optimal policy k -RPI or even analyzing multi-thresholds strategies are exciting future questions, even for $k = 1$.

A natural extension is to explore k -RPI under natural combinatorial constraints such as cardinality or matroid constraints, where the gambler can select multiple values while satisfying feasibility conditions, as it has been studied for the classical prophet inequality [16].

The k -RPI problem is very pessimistic as the k largest random variables are removed from the observed sequence. A more relaxed model would consider probabilities of failure. For example, a possibility could be where the i -th largest variable is removed with probability p_i . This is related to the model by [15, 39, 40] where $p_i = p$ for all i .

Finally, an interesting extension is to study if better competitive ratios for k -RPI can be obtained when the removed variables are not necessarily the largest, and the gambler has some offline information regarding the variables and/or the values removed.

Acknowledgements. Dana Pizarro was funded by the Chilean Agency for Research and Development (ANID) through PIA/PUENTE AFB230002, and the Fondecyt project number 11240705.

Bruno Ziliotto was funded by the French Agence Nationale de la Recherche (ANR) under reference ANR-21-CE40-0020 (CONVERGENCE project), and by public grants as part of the *Investissements d'avenir program*, reference ANR-11-LABX-0056-LMH, LabEx LMH and ANR-17-EURE-0010.

References

- [1] Krengel, U., Sucheston, L.: Semiamarts and finite values. *Bulletin of the American Mathematical Society* **83**(4), 745–747 (1977)
- [2] Hill, T.P., Kertz, R.P.: Comparisons of stop rule and supremum expectations of iid random variables. *The Annals of Probability*, 336–345 (1982)
- [3] Samuel-Cahn, E.: Comparison of threshold stop rules and maximum for independent nonnegative random variables. *the Annals of Probability*, 1213–1216 (1984)
- [4] Chawla, S., Hartline, J.D., Malec, D.L., Sivan, B.: Multi-parameter mechanism design and sequential posted pricing. In: *Proceedings of the Forty-second ACM Symposium on Theory of Computing*, pp. 311–320 (2010)
- [5] Correa, J., Foncea, P., Pizarro, D., Verdugo, V.: From pricing to prophets, and back! *Operations Research Letters* **47**(1), 25–29 (2019)
- [6] Hajiaghayi, M.T., Kleinberg, R., Sandholm, T.: Automated online mechanism design and prophet inequalities. In: *AAAI*, vol. 7, pp. 58–65 (2007)
- [7] Gallego, G., Segev, D.: A constructive prophet inequality approach to the adaptive probemax problem. *arXiv preprint arXiv:2210.07556* (2022)

- [8] Goyal, V., Udwani, R.: Online matching with stochastic rewards: Optimal competitive ratio via path-based formulation. *Operations Research* **71**(2), 563–580 (2023)
- [9] Huang, Z., Zhang, Q.: Online primal dual meets online matching with stochastic rewards: configuration lp to the rescue. In: *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pp. 1153–1164 (2020)
- [10] Vanderbei, R.J.: The postdoc variant of the secretary problem. Technical report, Tech. Report (2012)
- [11] Rose, J.S.: A problem of optimal choice and assignment. *Operations Research* **30**(1), 172–181 (1982)
- [12] Abels, A., Pitschmann, E., Schmand, D.: Prophet inequalities over time. In: *Proceedings of the 24th ACM Conference on Economics and Computation*, pp. 1–20 (2023)
- [13] Arsenis, M., Kleinberg, R.: Individual fairness in prophet inequalities. In: *Proceedings of the 23rd ACM Conference on Economics and Computation*, pp. 245–245 (2022)
- [14] Disser, Y., Fearnley, J., Gairing, M., Göbel, O., Klimm, M., Schmand, D., Skopalik, A., Tönnis, A.: Hiring secretaries over time: The benefit of concurrent employment. *Mathematics of Operations Research* **45**(1), 323–352 (2020)
- [15] Perez-Salazar, S., Singh, M., Toriello, A.: Robust online selection with uncertain offer acceptance. *Mathematics of Operations Research* (2024)
- [16] Kleinberg, R., Weinberg, S.M.: Matroid prophet inequalities. In: *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, pp. 123–136 (2012)
- [17] Rinott, Y., Samuel-Cahn, E.: Comparisons of optimal stopping values and prophet inequalities for negatively dependent random variables. *The Annals of Statistics* **15**(4), 1482–1490 (1987)
- [18] Rinott, Y., Samuel-Cahn, E.: Orderings of optimal stopping values and prophet inequalities for certain multivariate distributions. *Journal of multivariate analysis* **37**(1), 104–114 (1991)
- [19] Rinott, Y., Samuel-Cahn, E.: Optimal stopping values and prophet inequalities for some dependent random variables. *Lecture Notes-Monograph Series*, 343–358 (1992)
- [20] Rubinstein, A., Wang, J.Z., Weinberg, S.M.: Optimal single-choice prophet inequalities from samples. In: *11th Innovations in Theoretical Computer Science*

Conference (2020)

- [21] Lindley, D.V.: Dynamic programming and decision theory. *J. Roy. Statist. Soc. Ser. C Appl. Statist.* **10**, 39–52 (1961)
- [22] Dynkin, E.B.: The optimum choice of the instant for stopping a markov process. *Soviet Math. Dokl.* **4**, 627–629 (1963)
- [23] Ferguson, T.S.: Who solved the secretary problem? *Statist. Sci.* **4**, 282–296 (1989)
- [24] Gilbert, J.P., Mosteller, F.: Recognizing the maximum of a sequence. *Journal of the American Statistical Association* **61**(313), 35–73 (1966)
- [25] Correa, J., Foncea, P., Hoeksma, R., Oosterwijk, T., Vredeveld, T.: Posted price mechanisms and optimal threshold strategies for random arrivals. *Mathematics of operations research* **46**(4), 1452–1478 (2021)
- [26] Alaei, S.: Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. *SIAM Journal on Computing* **43**(2), 930–972 (2014)
- [27] Chawla, S., Hartline, J.D., Malec, D.L., Sivan, B.: Multi-parameter mechanism design and sequential posted pricing. In: *STOC 2010*, pp. 311–320 (2010)
- [28] Dütting, P., Feldman, M., Kesselheim, T., Lucier, B.: Prophet inequalities made easy: Stochastic optimization by pricing nonstochastic inputs. *SIAM Journal on Computing* **49**(3), 540–582 (2020)
- [29] Hajiaghayi, M.T., Kleinberg, R., Sandholm, T.: Automated online mechanism design and prophet inequalities. In: *AAAI 2007*, pp. 58–65 (2007)
- [30] Kleinberg, R., Weinberg, S.M.: Matroid prophet inequalities. In: *STOC 2012*, pp. 123–136 (2012)
- [31] Correa, J., Cristi, A., Epstein, B., Soto, J.: The two-sided game of googol. *Journal of Machine Learning Research* **23**(113), 1–37 (2022)
- [32] Nuti, P., Vondrák, J.: Secretary problems: The power of a single sample. In: *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 2015–2029 (2023). SIAM
- [33] Ezra, T., Feldman, M., Kupfer, R.: Prophet inequality with competing agents. In: *Algorithmic Game Theory: 14th International Symposium, SAGT 2021, Aarhus, Denmark, September 21–24, 2021, Proceedings 14*, pp. 112–123 (2021). Springer
- [34] Gensbittel, F., Pizarro, D., Renault, J.: Competition and recall in selection problems. *Dynamic Games and Applications* **14**(4), 806–845 (2024)
- [35] Immorlica, N., Kleinberg, R., Mahdian, M.: Secretary problems with competing

- employers. In: International Workshop on Internet and Network Economics, pp. 389–400 (2006). Springer
- [36] Karlin, A., Lei, E.: On a competitive secretary problem. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 29 (2015)
- [37] Ramsey, D.: A stackelberg game based on the secretary problem: Optimal response is history dependent. arXiv preprint arXiv:2409.04153 (2024)
- [38] Perez-Salazar, S., Verdugo, V.: Optimal guarantees for online selection over time. arXiv preprint arXiv:2408.11224 (2024)
- [39] Smith, M.: A secretary problem with uncertain employment. Journal of applied probability **12**(3), 620–624 (1975)
- [40] Tamaki, M.: A secretary problem with uncertain employment and best choice of available candidates. Operations Research **39**(2), 274–284 (1991)
- [41] Perez-Salazar, S., Singh, M., Toriello, A.: The iid prophet inequality with limited flexibility. Mathematics of Operations Research (2025)

Appendix A Random Order RPI

In this subsection, we justify that in any information model for k -RPI, if the gambler observes the values in random order, then there is an algorithm that attains a competitive ratio of at least $1/e$. The result follows from the following,

Lemma 4. *In the NI k -RPI, the secretary problem algorithm guarantees a competitive ratio of at least $1/e$ if the values are observed in a random order.*

Proof. Let X_1, \dots, X_n be the random values and let $X_{(k+1)} \geq \dots \geq X_{(n)}$ be the $n - k$ values observed by the gambler. Applying the secretary problem algorithm, guarantees that the gambler accepts $X_{(k+1)}$ with probability at least $1/e$. From here, the result follows. \square

Appendix B Missing Proofs From Section 3

B.1 Proof of Lemma 1

The proof of Lemma 1 is divided into two intermediary results, that we state now.

Lemma 5.

$$\mathbb{E}(\overline{MSA}_{k+1} 1_{\tau_{k+1}=w_{2k+2}}) \geq \sum_{l=k+1}^{2k+1} \mathbb{P}(X_{(k+1)} = w_l) w_l \frac{\delta_l}{2}.$$

Lemma 6. Assume that $2k + 2 \neq \xi_{k+1}$. Then

$$\mathbb{E}(\overline{MSA}_{k+1} 1_{\tau_{k+1} \leq w_{2k+3}}) \geq \frac{1}{2} \sum_{l=2(k+1)}^{\xi_{k+1}} \mathbb{P}(X_{(k+1)} = w_l) w_l.$$

In the case where $2k + 2 \neq \xi_{k+1}$, summing the two inequalities proves Lemma 1. Assume that $2k + 2 = \xi_{k+1}$. This means that the elements of $\{w_1, \dots, w_{2k+2}\}$ form $k+1$ pairs. Hence, if $w_{2k+2} \in S$, which happens with probability $1/2$, there are exactly $k+1$ elements larger than w_{2k+2} that are in X . In that case, \overline{MSA}_{k+1} picks $X_{(k+1)}$. It follows that

$$\mathbb{E}(\overline{MSA}_{k+1}) \geq \frac{1}{2} \mathbb{E}(X_{(k+1)}).$$

In particular, Lemma 1 holds. We dedicate one sub-section to each lemma.

B.1.1 Proof of Lemma 5

Lemma 5 is a consequence of the following lemma.

Lemma 7. Let $l \in \{1, \dots, 2k+1\}$ such that w_l is not blocked.

a) If w_l is not ill-paired, it holds that

$$\mathbb{P}(\overline{MSA}_{k+1} = w_l \cap \tau_{k+1} = w_{2k+2} | X_{(k+1)} = w_l) \geq 2^{-2k-2+l}.$$

b) If w_l is ill-paired, then

$$\mathbb{P}(\overline{MSA}_{k+1} = w_l \cap \tau_{k+1} = w_{2k+2} | X_{(k+1)} = w_l) \geq 2^{-2k-1+l}$$

Proof of Lemma 7. a) We claim that when $X_{(k+1)} = w_l$ and $\tau_{k+1} = w_{2k+2}$, then \overline{MSA}_{k+1} picks w_l . Indeed, when $X_{(k+1)} = w_l$, there are exactly $l-1-k$ elements in $\{w_1, \dots, w_{l-1}\}$ that are in S . If, in addition, $\tau_{k+1} = w_{2k+2}$, then there should be exactly $k - (l-1-k) = 2k+1-l$ elements of $\{w_{l+1}, \dots, w_{2k+1}\}$ that are in S , meaning that they should all be in S . Under these circumstances, w_l is the only element in X that is above τ_{2k+2} and that is not among the k best values in X , and is thus selected by \overline{MSA}_{k+1} . We deduce that

$$\begin{aligned} & \mathbb{P}(\{\overline{MSA}_{k+1} = w_l\} \cap \{\tau_{k+1} = w_{2k+2}\} | X_{(k+1)} = w_l) \\ &= \mathbb{P}(\tau_{k+1} = w_{2k+2} | X_{(k+1)} = w_l). \end{aligned}$$

Therefore, it is enough to prove $\mathbb{P}(\tau_{k+1} = w_{2k+2} | X_{(k+1)} = w_l) \geq 2^{-2k-2+l}$. Given $X_{(k+1)} = w_l$, in order for $\tau_{k+1} = w_{2k+2}$ to hold, it is necessary and sufficient that all the elements in $\{w_{l+1}, \dots, w_{2k+2}\}$ belong to S . We claim that this event occurs with probability greater than $2^{-2k+l-2}$. To show that, we use the chain rule for conditional probability:

$$\begin{aligned}\mathbb{P}(\{w_{l+1}, \dots, w_{2k+2}\} \subset S | X_{k+1} = w_l) &= \mathbb{P}\left(\bigcap_{j=l+1}^{2k+2} \{w_j \in S\} \mid X_{(k+1)} = w_l\right) \\ &= \prod_{l'=l+1}^{2k+2} \mathbb{P}\left(\{w_{l'} \in S\} \mid \bigcap_{j=l+1}^{l'-1} \{w_j \in S\}, X_{(k+1)} = w_l\right).\end{aligned}$$

In order to establish the desired result, it is sufficient to verify that each factor in the expression above is lower bounded by $1/2$. The proof is therefore divided into two steps:

Step 1: $\mathbb{P}(w_{l+1} \in S | X_{(k+1)} = w_l) \geq 1/2$.

If w_{l+1} is paired with an element smaller than w_{l+1} , then the events $\{w_{l+1} \in S\}$ and $\{X_{(k+1)} = w_l\}$ are independent, and therefore

$$\mathbb{P}(w_{l+1} \in S | X_{(k+1)} = w_l) = 1/2.$$

Consider now the case where w_{l+1} is paired with some $w_a \geq w_{l+1}$. Since w_l is not ill-paired, we have $a \neq l$, and the probability that w_{l+1} lies in S is equal to the probability that w_a is one of the not-paired elements of $\{w_1, \dots, w_l\}$ in X . Note that the event $\{X_{(k+1)} = w_l\}$ occurs if and only if $w_l \in X$ and there are exactly k elements in X that are larger than w_l . Therefore, if $l \in \{\xi_j, \dots, \xi_{j+1} - 1\}$, then among the $l - 2j$ not-paired values in $\{w_1, \dots, w_l\}$, $k + 1 - j$ belong to X , while $l - k - 1 - j$ are in S . It follows that the probability of w_a being among those elements in X is higher than $1/2$, since $k + 1 - j > l - k - 1 - j$ due to $l \leq 2k + 1$. We thus conclude that

$$\mathbb{P}(w_{l+1} \in S | X_{k+1} = w_l) > \frac{1}{2}.$$

Step 2: For each $l' \in \{l + 2, \dots, 2k + 2\}$,

$$\mathbb{P}\left(w_{l'} \in S \mid \bigcap_{j=l+1}^{l'-1} \{w_j \in S\}, X_{(k+1)} = w_l\right) \geq \frac{1}{2}.$$

Let w_a such that $w_{l'}$ is paired with w_a , and assume that $l' \in \{\xi_j + 1, \dots, \xi_{j+1}\}$. That is, there are j pairs that arrived before $w_{l'}$. Following the same argument than in Step 1, if $w_a < w_{l'}$, we have

$$\mathbb{P}\left(w_{l'} \in S \mid \bigcap_{j=l+1}^{l'-1} \{w_j \in S\}, X_{(k+1)} = w_l\right) = \frac{1}{2}.$$

On the other hand, w_a cannot belong to $\{w_{l+1}, \dots, w_{l'-1}\}$ because w_l is not blocked.

Finally, let us assume $w_a \leq w_l$. In this case, among the $l' - 1 - 2j$ not-paired values in $\{w_1, \dots, w_{l'-1}\}$, $k+1-j$ belong to X , while $l' - 2 - j - k$ are in S . Then, it follows that the probability of w_a being among those elements in X is higher than $1/2$, since $k+1-j > l' - 2 - j - k$ due to $l' \leq 2k+1$. We thus conclude that

$$\mathbb{P}\left(w_{l'} \in S \mid \bigcap_{j=l+1}^{l'-1} \{w_j \in S\}, X_{(k+1)} = w_l\right) > \frac{1}{2}.$$

Combining Step 1 and Step 2 yields the result.

b) As in the proof of Case a), we have

$$\mathbb{P}(\overline{MSA}_{k+1} = w_l \cap \{\tau_{k+1} = w_{2k+2}\} \mid X_{(k+1)} = w_l) = \mathbb{P}(\tau_{k+1} = w_{2k+2} \mid X_{(k+1)} = w_l),$$

and to obtain the result it is enough to show that

$$\mathbb{P}(\tau_{k+1} = w_{2k+2} \mid X_{(k+1)} = w_l) \geq 2^{-2k-1+l}. \quad (\text{B1})$$

Given that $X_{(k+1)} = w_l$, we know that w_p is in S , since (w_p, w_l) is a pair. Then, in order to get $\tau_{k+1} = w_{2k+2}$, it is necessary and sufficient that all the elements in $\{w_{l+1}, \dots, w_{2k+2}\} \setminus \{p\}$ belong to S . This happens with probability at least $2^{-2k-1+l}$, by the same argument as in the proof of Case a). This proves (B1), and the result follows. \square

B.1.2 Proof of Lemma 6

First, we decompose the left-hand-side term in Lemma 5 as follows:

$$\mathbb{E}(\overline{MSA}_{k+1} 1_{\tau_{k+1} \leq w_{2k+3}}) = \sum_{l=2k+3}^{\xi_{k+1}} \mathbb{E}(\overline{MSA}_{k+1} \mid \tau_{k+1} = w_l) \mathbb{P}(\tau_{k+1} = w_l)$$

Since each w_l , for $l \geq 1$, is equally likely to be in S or in X , the law of τ_{k+1} is identical to the law of $X_{(k+1)}$. We deduce that for all $l \geq 1$, $\mathbb{P}(\tau_{k+1} = w_l) = \mathbb{P}(X_{(k+1)} = w_l)$. Secondly, when $\tau_{k+1} = w_l$, there are $l - k - 1 \geq k+1$ elements above w_l that are in X . Hence, \overline{MSA}_{k+1} will pick one of them, and we deduce that $\mathbb{E}(\overline{MSA}_{k+1} \mid \tau_{k+1} = w_l) \geq w_{l-1}$. These two observations give

$$\mathbb{E}(\overline{MSA}_{k+1} 1_{\tau_{k+1} \leq w_{2k+3}}) \geq \sum_{l=2k+3}^{\xi_{k+1}} w_{l-1} \mathbb{P}(X_{k+1} = w_l)$$

$$= \sum_{l=2(k+1)}^{\xi_{k+1}-1} w_l \mathbb{P}(X_{(k+1)} = w_{l+1}) \quad (\text{B2})$$

One of the main differences between the above inequality and the one we want to prove in Lemma 6 is that the term inside the sum is $\mathbb{P}(X_{(k+1)} = w_{l+1})$ instead of $\mathbb{P}(X_{(k+1)} = w_l)$. In the sequel, we relate these two quantities. First, we compute $\mathbb{P}(X_{(k+1)} = w_l)$.

Lemma 8. *The probability distribution of $X_{(k+1)}$ is given by*

$$\mathbb{P}(X_{(k+1)} = w_l) = \begin{cases} \frac{\binom{l-1-2j}{k-j}}{2^{l-2j}} & \text{if } l \in \{\xi_j + 1, \dots, \xi_{j+1} - 1\}, \text{ for } j \in \{0, \dots, k\} \\ \frac{\binom{\xi_j-2j}{k+1-j}}{2^{\xi_j-2j+1}} & \text{if } l = \xi_j, \text{ for } j \in \{1, \dots, k+1\}. \end{cases}$$

Proof of Lemma 8. We divide the proof into two cases, depending on whether $l = \xi_j$ for some $j \in \{0, \dots, k+1\}$ or not.

Case 1: Suppose that $l \in \{\xi_j + 1, \dots, \xi_{j+1} - 1\}$ for some $j \in \{0, \dots, k\}$. Note that $X_{(k+1)} = w_l$ if and only if $w_l \in X$ and there are exactly k values in X that are larger than w_l .

Since $l \in \{\xi_j + 1, \dots, \xi_{j+1} - 1\}$, we have, conditioned on $w_l \in X$, that there are $j+1$ values in X and j in S with probability 1. Therefore, the probability that exactly k values in X are among the $l-1$ largest values is given by

$$\binom{l-2j-1}{k-j} \frac{1}{2^{l-2j-1}}.$$

On the other hand, $\mathbb{P}(w_l \in X) = 1/2$, and thus we conclude that in this case,

$$\mathbb{P}(X_{(k+1)} = w_l) = \binom{l-2j-1}{k-j} \frac{1}{2^{l-2j}}. \quad (\text{B3})$$

Case 2: Suppose that $l = \xi_j$ for some $j \in \{1, \dots, k+1\}$. The analysis in this case is similar to that of Case 1. However, the probability of having exactly k values in X greater than w_l is now given by

$$\binom{l-2j}{k-(j-1)} \frac{1}{2^{l-2j}}.$$

In effect, conditioned on $w_l \in X$, there are $j-1$ values in X greater than w_l and j values greater than w_l in S , with probability one. Thus, we need to compute the probability that exactly $k-(j-1)$ additional values in X come from the $l-2j$ remaining elements. This probability is given by the expression above.

Therefore, in this case,

$$\mathbb{P}(X_{(k+1)} = w_l) = \binom{l-2j}{k-(j-1)} \frac{1}{2^{l-2j+1}} \quad (\text{B4})$$

Combining (B3) and (B4), we obtain the desired result. \square

We now use the previous lemma to lower bound $\mathbb{P}(X_{(k+1)} = w_{l+1})$ in terms of $\mathbb{P}(X_{(k+1)} = w_l)$. As suggested by the expression in Lemma 8, we will need to distinguish between the cases where l and $l+1$ are some ξ_j or not.

Lemma 9. *a) Let $j \in \{0, \dots, k\}$ and $l \in \{\xi_j + 1, \dots, \xi_{j+1} - 1\}$.*

$$\mathbb{P}(X_{(k+1)} = w_{l+1}) \geq \begin{cases} \frac{1}{2} \mathbb{P}(X_{(k+1)} = w_l) & \text{if } l+1 \in \{\xi_j + 1, \dots, \xi_{j+1} - 1\} \\ \mathbb{P}(X_{(k+1)} = w_l) & \text{if } l+1 = \xi_{j+1}. \end{cases}$$

b) Assume that $2k+2 = \xi_j$, for some $j \in \{1, \dots, k\}$. Then

$$\mathbb{P}(X_{(k+1)} = w_{2k+3}) \geq \begin{cases} \frac{1}{2} \mathbb{P}(X_{(k+1)} = w_{2k+2}) & \text{if } 2k+3 \neq \xi_{j+1} \\ \mathbb{P}(X_{(k+1)} = w_{2k+2}) & \text{if } 2k+3 = \xi_{j+1}. \end{cases}$$

Proof of Lemma 9. *a) Assume $l+1 \in \{\xi_j + 1, \dots, \xi_{j+1} - 1\}$. We have*

$$\begin{aligned} \mathbb{P}(X_{(k+1)} = w_{l+1}) &= \frac{\binom{l-2j}{k-j}}{2^{l+1-2j}} \\ &\geq \frac{1}{2} \cdot \frac{\binom{l-1-2j}{k-j}}{2^{l-2j}} \\ &= \frac{1}{2} \cdot \mathbb{P}(X_{(k+1)} = w_l) \end{aligned}$$

Assume that $l+1 = \xi_{j+1}$. We have

$$\begin{aligned} \mathbb{P}(X_{(k+1)} = w_{l+1}) &= \frac{\binom{\xi_{j+1}-2j-2}{k-j}}{2^{\xi_{j+1}-2j-1}} \\ &= \mathbb{P}(X_{(k+1)} = w_l) \end{aligned}$$

b) Assume that $2k+3 \neq \xi_{j+1}$. We have

$$\begin{aligned} \mathbb{P}(X_{(k+1)} = w_{2k+3}) &= \frac{\binom{2k+2-2j}{k-j}}{2^{2k+3-2j}} \\ &= \left(\frac{k+1-j}{k+2-j} \right) \frac{\binom{2k+2-2j}{k+1-j}}{2^{2k+3-2j}} \end{aligned}$$

$$\geq \frac{1}{2} \mathbb{P}(X_{(k+1)} = w_{2k+2})$$

Assume that $2k+3 = \xi_{j+1}$. We have

$$\begin{aligned} \mathbb{P}(X_{(k+1)} = w_{2k+3}) &= \frac{\binom{2k+1-2j}{k-j}}{2^{2k+2-2j}} \\ &= \frac{\frac{1}{2} \binom{2k+2-2j}{k+1-j}}{2^{2k+2-2j}} \\ &= \mathbb{P}(X_{(k+1)} = w_{2k+2}) \end{aligned}$$

□

We are now ready to prove Lemma 6.

Proof of Lemma 6. By inequality (B2), it is enough to prove that

$$\sum_{l=2(k+1)}^{\xi_{k+1}-1} w_l \mathbb{P}(X_{(k+1)} = w_{l+1}) \geq \frac{1}{2} \sum_{l=2k+2}^{\xi_{k+1}} w_l \mathbb{P}(X_{(k+1)} = w_l).$$

Case 1. $2k+3 = \xi_j$, for some $j \in \{1, \dots, k\}$.

By Lemma 9 b), we have

$$\mathbb{P}(X_{(k+1)} = w_{2k+3}) \geq \frac{1}{2} \mathbb{P}(X_{(k+1)} = w_{2k+2}) + \frac{1}{2} \mathbb{P}(X_{(k+1)} = w_{2k+3}) \quad (\text{B5})$$

We deduce that

$$\begin{aligned} \sum_{l=2(k+1)}^{\xi_{k+1}-1} w_l \mathbb{P}(X_{(k+1)} = w_{l+1}) &= w_{2k+2} \mathbb{P}(X_{(k+1)} = w_{2k+3}) + \sum_{l=2k+3}^{\xi_{k+1}-1} w_l \mathbb{P}(X_{(k+1)} = w_{l+1}) \\ &\geq \frac{1}{2} w_{2k+2} \mathbb{P}(X_{(k+1)} = w_{2k+2}) + \frac{1}{2} w_{2k+3} \mathbb{P}(X_{(k+1)} = w_{2k+3}) \\ &\quad + \sum_{l=2k+3}^{\xi_{k+1}-1} w_{l+1} \mathbb{P}(X_{(k+1)} = w_{l+1}) \\ &\geq \frac{1}{2} \sum_{l=2k+2}^{\xi_{k+1}} w_l \mathbb{P}(X_{(k+1)} = w_l), \end{aligned}$$

where in the second-to-last inequality, we used (B5) and the fact that $w_{2k+2} \geq w_{2k+3}$ and $w_l \geq w_{l+1}$.

Case 2. $2k+3 \in \{\xi_j + 1, \dots, \xi_{j+1} - 1\}$ for some $j \in \{0, \dots, k\}$.

The sum $\sum_{l=2k+2}^{\xi_{k+1}-1} w_l \mathbb{P}(X_{(k+1)} = w_{l+1})$ can be decomposed as

$$\sum_{l=2k+2}^{\xi_{j+1}-2} w_l \mathbb{P}(X_{(k+1)} = w_{l+1}) + w_{\xi_{j+1}-1} \mathbb{P}(X_{(k+1)} = w_{\xi_{j+1}}) + \sum_{l=\xi_{j+1}}^{\xi_{k+1}-1} w_l \mathbb{P}(X_{(k+1)} = w_{l+1}).$$

By using Lemma 9 and the fact that $w_l \geq w_{l+1}$, we can lower bound the expression above by

$$\frac{1}{2} \sum_{l=2k+2}^{\xi_{j+1}-2} w_l \mathbb{P}(X_{(k+1)} = w_l) + w_{\xi_{j+1}-1} \mathbb{P}(X_{(k+1)} = w_{\xi_{j+1}-1}) + \sum_{l=\xi_{j+1}}^{\xi_{k+1}-1} w_{l+1} \mathbb{P}(X_{(k+1)} = w_{l+1}),$$

which is at least

$$\frac{1}{2} \sum_{l=2k+2}^{\xi_{k+1}} w_l \mathbb{P}(X_{(k+1)} = w_l),$$

as we wanted to see. □

B.2 Proof of Lemma 2

Before proving Lemma 2, we introduce one technical lemma that gives a lower bound for the probability of \overline{MSA}_i picking a value w_l conditional on w_l being the $(k+1)$ -largest value in the set X .

Lemma 10. *Let $l \in \{1, \dots, 2k+1\}$.*

a) If w_l is not blocked and not ill-paired, for all $i \in \{l-k, \dots, k\}$ it holds

$$\mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) \geq 2^{-k-i+l-1}.$$

b) If w_l is blocked, and that either w_l is not ill-paired, or it is ill-paired and $m_l < p$, we have

$$\mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) \geq \begin{cases} 2^{-k-i+l-1} & \text{if } i \in \{l-k, \dots, m_l-k-3\}, \\ 2^{-k-i+l} & \text{if } i = m_l-k-2. \end{cases}$$

c) If w_l is not blocked and ill-paired. Then,

$$\mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) \geq \begin{cases} 2^{-k-i+l-1} & \text{if } i \in \{l-k, \dots, p-k-2\}, \\ 2^{-k-i+l} & \text{if } i \in \{p-k, \dots, k\}. \end{cases}$$

d) If w_l is blocked and ill-paired, and that $m_l > p$. Then,

$$\mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) \geq \begin{cases} 2^{-k-i+l-1} & \text{if } i \in \{l-k, \dots, p-k-2\}, \\ 2^{-k-i+l} & \text{if } i \in \{p-k, \dots, m_l-k-3\}, \\ 2^{-k-i+l+1} & \text{if } i = m_l-k-2. \end{cases}$$

Proof of Lemma 10. Let $l \in \{1, \dots, 2k+1\}$.

a) The proof is very similar to the one of Lemma 7 a), up to replacing $k+1$ by i . For sake of completeness, we draw the main lines. Take $i \in \{l-k, \dots, k\}$. We want to analyze $\mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l)$. First, note that since l is not ill-paired, w_l is not paired with w_{k+i+1} . Then,

$$\begin{aligned} \mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) &\geq \mathbb{P}(\{\overline{MSA}_i = w_l\} \cap \{\tau_i = w_{k+i+1}\} | X_{(k+1)} = w_l) \\ &= \mathbb{P}(\tau_i = w_{k+i+1} | X_{(k+1)} = w_l), \end{aligned}$$

where the equality stems from the fact that, when $X_{(k+1)} = w_l$ and $\tau_i = w_{k+i+1}$, all the elements in $\{w_{l+1}, \dots, w_{k+i}\}$ must be in S , and then \overline{MSA}_i picks w_l , because it is not paired with the threshold τ_i .

Therefore, it is enough to prove that $\mathbb{P}(\tau_i = w_{k+i+1} | X_{(k+1)} = w_l) \geq 2^{-k-i+l-1}$. Given $X_{(k+1)} = w_l$, in order for $\tau_i = w_{k+i+1}$ to hold, it is necessary and sufficient that all the elements in $\{w_{l+1}, \dots, w_{k+i+1}\}$ belong to S . This event occurs with probability greater than $2^{-k-i+l-1}$, by a similar computation as in the proof of Lemma 7 a).

b) Take $i \in \{l-k, \dots, m_l-k-3\}$. In this case, $k+i+1 < m_l$, hence the pair that blocks w_l is smaller than w_{k+i+1} . Moreover, since either w_l is not ill-paired or $m_l < p$, w_l is not paired with w_{k+i+1} . We can therefore replicate the same computations as in a), and thus obtain the claimed inequality.

If $i = m_l - k - 2$, we can replicate the same computations as in a) too, which yields:

$$\mathbb{P}(\overline{MSA}_i = w_l \cap \{\tau_i = w_{k+i+1}\} | X_{(k+1)} = w_l) \geq 2^{-k-i+l-1}.$$

To obtain the desired lower bound, we will consider in addition the case where $\tau_i = w_{m_l}$. Indeed, whenever $X_{(k+1)} = w_l$ and $\tau_i = w_{m_l}$, the only element in X that is below w_l and above the threshold τ_i is w_{m_l} 's pair, namely $w_{m'}$. By definition of \overline{MSA}_i , such an element is not selected, and therefore \overline{MSA}_i selects w_l . We deduce that

$$\mathbb{P}(\{\overline{MSA}_i = w_l\} \cap \{\tau_i = w_{m_l}\} | X_{(k+1)} = w_l) = \mathbb{P}(\tau_i = w_{m_l} | X_{(k+1)} = w_l)$$

Knowing $X_{(k+1)} = w_l$, in order to get $\tau_i = w_{m_l} = w_{k+i+2}$, it is necessary and sufficient that all the elements in $\{w_{l+1}, \dots, w_{k+i+2}\} \setminus \{m'\}$ belong to S , which

happens with probability higher than $2^{-k-i+l-1}$, by a similar computation as in the proof of Lemma 7 a). Then, we have

$$\mathbb{P}(\overline{MSA}_i = w_l \cap \{\tau_i = w_{k+i+2}\} | X_{(k+1)} = w_l) \geq 2^{-k-i+l-1}.$$

We conclude that

$$\begin{aligned} \mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) &\geq \mathbb{P}(\overline{MSA}_i = w_l \cap \{\tau_i = w_{k+i+2}\} | X_{(k+1)} = w_l) \\ &\quad + \mathbb{P}(\overline{MSA}_i = w_l \cap \{\tau_i = w_{k+i+1}\} | X_{(k+1)} = w_l) \\ &\geq 2^{-k-i+l}, \end{aligned}$$

which is the desired result.

c) If $i \in \{l-k, \dots, p-k-2\}$, the argument proceeds as in part a).

Take $i \in \{p-k, \dots, k\}$. In this case, $k+i+1 > p$, and then w_l is not paired with w_{k+i+1} . We therefore have

$$\begin{aligned} \mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) &\geq \mathbb{P}(\overline{MSA}_i = w_l \cap \{\tau_i = w_{k+i+1}\} | X_{(k+1)} = w_l) \\ &= \mathbb{P}(\tau_i = w_{k+i+1} | X_{(k+1)} = w_l), \end{aligned}$$

and to obtain the result it is enough to show that

$$\mathbb{P}(\tau_i = w_{k+i+1} | X_{(k+1)} = w_l) \geq 2^{-k-i+l}.$$

Given that $X_{(k+1)} = w_l$, we know that w_p is in S , since (w_p, w_l) is a pair. Then, in order to get $\tau_i = w_{k+i+1}$, it is necessary and sufficient that all the elements in $\{w_{l+1}, \dots, w_{k+i+1}\} \setminus \{p\}$ belong to S , which happens with probability at least 2^{-k-i+l} , by a similar computation as in the proof of Lemma 7 a). We deduce that

$$\mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) \geq 2^{-k-i+l},$$

which is what we wanted to show.

d) The first two cases can be proved as in a) and c). Let $i = m_l - k - 2$; that is $k+i+2 = m_l$. We call $w_{m'}$ to the pair of w_{m_l} .

First, we can replicate the computations of Case c), and obtain:

$$\mathbb{P}(\overline{MSA}_i = w_l \cap \{\tau_i = w_{k+i+1}\} | X_{(k+1)} = w_l) \geq 2^{-k-i+l}.$$

As in Case b), in order to obtain the claimed bound of the lemma, we need to consider the event $\{\tau_i = w_{k+i+2}\}$. We have

$$\begin{aligned} \mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) &\geq \mathbb{P}(\overline{MSA}_i = w_l \cap \{\tau_i = w_{k+i+2}\} | X_{(k+1)} = w_l) \\ &= \mathbb{P}(\tau_i = w_{k+i+2} | X_{(k+1)} = w_l), \end{aligned}$$

where the equality stems from the fact that, when $X_{(k+1)} = w_l$ and $\tau_i = w_{k+i+2}$, all the elements in $\{w_{l+1}, \dots, w_{k+i+1}\} \setminus \{w_{m'}\}$ must be in S and then \overline{MSA}_i picks w_l , because w_l is not paired with w_{k+i+2} . Since w_l and w_p are paired, given that $X_{(k+1)} = w_l$, we have that w_p lies in S . Moreover, since w_{k+i+2} and $w_{m'}$ are paired, if w_{k+i+2} lies in S , then $w_{m'}$ lies in X . Hence, in order to get $\tau_i = w_{k+i+2}$, it is necessary and sufficient that all the elements in $\{w_{l+1}, \dots, w_{k+i+2}\} \setminus \{w_{m'}, w_p\}$ belong to S , which happens with probability at least 2^{-k-i+l} , by a similar computation as in the proof of Lemma 7 a). We deduce that

$$\mathbb{P}(\overline{MSA}_i = w_l \cap \{\tau_i = w_{k+i+2}\} | X_{(k+1)} = w_l) \geq 2^{-k-i+l}.$$

We conclude that

$$\begin{aligned} \mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) &\geq \mathbb{P}(\overline{MSA}_i = w_l \cap \{\tau_i = w_{k+i+2}\} | X_{(k+1)} = w_l) \\ &\quad + \mathbb{P}(\overline{MSA}_i = w_l \cap \{\tau_i = w_{k+i+1}\} | X_{(k+1)} = w_l) \\ &\geq 2^{-k-i+l+1}. \end{aligned}$$

□

Proof of Lema 2. Note that

$$\begin{aligned} \sum_{i=1}^k \mathbb{E}(\overline{MSA}_i) &\geq \sum_{i=1}^k \sum_{l=1}^{2k+1} \mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) w_l \mathbb{P}(X_{(k+1)} = w_l) \\ &= \sum_{l=1}^{2k+1} w_l \mathbb{P}(X_{(k+1)} = w_l) \sum_{i=1}^k \mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l). \end{aligned}$$

In the remainder of the proof we show that for each $l \in \{1, \dots, 2k+1\}$,

$$\sum_{i=1}^k \mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) = 1 - \delta_l,$$

where

$$\delta_l = \begin{cases} 2^{-2k+l-1} & \text{if } w_l \text{ is not blocked and not ill-paired} \\ 2^{-2k+l} & \text{if } w_l \text{ is not blocked and ill-paired} \\ 0 & \text{otherwise.} \end{cases}$$

Case 1. w_l is not blocked and not ill-paired.

In this case, by Lemma 10

$$\sum_{i=1}^k \mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) = \sum_{i=l-k}^k 2^{-k-i+l-1} = 1 - 2^{-2k+l-1}.$$

Case 2. w_l is not blocked and ill-paired.

$$\sum_{i=1}^k \mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) = \sum_{i=l-k}^{p-k-2} 2^{-k-i+l-1} + \sum_{i=p-k}^k 2^{-k-i+l} = 1 - 2^{-2k+l}.$$

Case 3. w_l is blocked, and that either w_l is not ill-paired, or it is ill-paired and $m_l < p$.

In this case,

$$\sum_{i=1}^k \mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) = \sum_{i=l-k}^{m_l-k-3} 2^{-k-i+l-1} + 2^{l-m_l+2} = 1.$$

Case 4. w_l is blocked and ill-paired, and that $m_l > p$. In this case,

$$\sum_{i=1}^k \mathbb{P}(\overline{MSA}_i = w_l | X_{(k+1)} = w_l) = \sum_{i=l-k}^{p-k-2} 2^{-k-i+l-1} + \sum_{i=p-k}^{m_l-k-3} 2^{-k-i+l} + 2^{-m_l+l+3} = 1.$$

Putting everything together, we obtain the desired result. \square

B.3 Proof of Proposition 1

We want to prove that $\mathbb{E}(MSA_{RAND}) \geq \frac{1}{2(k+1)} \mathbb{E}(X_{(k+1)})$.

First, note that

$$\begin{aligned} (k+1)\mathbb{E}(MSA_{RAND}) &= \sum_{i=1}^{k+1} \mathbb{E}(MSA_i) \\ &= \sum_{i=1}^k \mathbb{E}(MSA_i | X_{(k+1)} = w_{k+i}) \mathbb{P}(X_{(k+1)} = w_{k+i}) + \mathbb{E}(MSA_{k+1}). \end{aligned}$$

Now, using that for each $i \in \{1, \dots, k\}$

$$\mathbb{E}(MSA_i | X_{(k+1)} = w_{k+i}) \geq w_{k+i} \mathbb{P}(MSA_i = w_{k+i} | X_{(k+1)} = w_{k+i}),$$

we have that $(k+1)\mathbb{E}(MSA_{RAND})$ is at least

$$\sum_{i=1}^k w_{k+i} \mathbb{P}(MSA_i = w_{k+i} | X_{(k+1)} = w_{k+i}) \mathbb{P}(X_{(k+1)} = w_{k+i}) + \mathbb{E}(MSA_{k+1}).$$

In the remainder of the proof we bound $\mathbb{P}(MSA_i = w_{k+i} | X_{(k+1)} = w_{k+i})$ for each $i \in \{1, \dots, k\}$, and $\mathbb{E}(MSA_{k+1})$.

Step 1. For each $i \in \{1, \dots, k\}$,

$$\mathbb{P}(MSA_i = w_{k+i} | X_{(k+1)} = w_{k+i}) \geq 1/2.$$

On one hand,

$$\begin{aligned} \mathbb{P}(MSA_i = w_{k+i} | X_{(k+1)} = w_{k+i}) &\geq \mathbb{P}(\{MSA_i = w_{k+i}\} \cap \{\tau_i = w_{k+i+1}\} | X_{(k+1)} = w_{k+i}) \\ &= \mathbb{P}(\tau_i = w_{k+i+1} | X_{(k+1)} = w_{k+i}), \end{aligned}$$

where the equality stems from the fact that, when $X_{(k+1)} = w_{k+i}$ and $\tau_i = w_{k+i+1}$, MSA_i picks w_{k+i} .

On the other hand, it is easy to see that $\mathbb{P}(\tau_i = w_{k+i+1} | X_{(k+1)} = w_l)$ is 1 if w_{k+i} and w_{k+i+1} are paired, and $1/2$, otherwise.

Therefore, we obtain $\mathbb{P}(\tau_i = w_{k+i+1} | X_{(k+1)} = w_l) \geq 2^{-1}$, and the first step is completed.

Step 2. $\mathbb{E}(MSA_{k+1}) \geq 1/2 \sum_{l=2k+1}^{\xi_{k+1}} w_l \mathbb{P}(X_{(k+1)} = w_l)$.

If $2k+2 = \xi_{k+1}$, the elements of $\{w_1, \dots, w_{2k+2}\}$ form $k+1$ pairs. Hence, if $w_{2k+2} \in S$, which happens with probability $1/2$, there are exactly $k+1$ elements larger than w_{2k+2} that are in X . In that case, MSA_{k+1} picks $X_{(k+1)}$. It follows that

$$\mathbb{E}(MSA_{k+1}) \geq \frac{1}{2} \mathbb{E}(X_{(k+1)}) \geq \frac{1}{2} \sum_{l=2k+1}^{\xi_{k+1}} w_l \mathbb{P}(X_{(k+1)} = w_l).$$

If $2k+2 \neq \xi_{k+1}$,

$$\mathbb{E}(MSA_{k+1}) = \mathbb{E}(MSA_{k+1} 1_{\tau_{k+1}=w_{2k+2}}) + \mathbb{E}(MSA_{k+1} 1_{\tau_{k+1} \leq w_{2k+3}}).$$

By Lemma 5,

$$\mathbb{E}(MSA_{k+1} 1_{\tau_{k+1}=w_{2k+2}}) \geq \mathbb{P}(X_{(k+1)} = w_{2k+1}) w_{2k+1} \frac{\delta_{2k+1}}{2},$$

where δ_{2k+1} is equal to 1 since by definition w_{2k+1} cannot be blocked nor ill-paired.

On the other hand, by Lemma 6, the second term is lower bounded by

$$\frac{1}{2} \sum_{l=2(k+1)}^{\xi_{k+1}} \mathbb{P}(X_{(k+1)} = w_l) w_l.$$

Putting all together, we obtain the step 2.

Combining Step 1 and Step 2, we conclude that

$$(k+1)\mathbb{E}(MSA_{RANDOM}) \geq \frac{1}{2} \sum_{l=k+1}^{\xi_{k+1}} w_l \mathbb{P}(X_{(k+1)} = w_l) = \mathbb{E}(X_{(k+1)}),$$

and the proof is completed.

Appendix C Missing Proofs From Section 5

C.1 Proof of Proposition 2

The probability of reaching $i+1$ is the same as not observing a value at least u among the first i trials, which is given by $\mathbb{P}(X_1, \dots, X_i < u \mid E)$, while the reward at $i+1$ is the expected value when $X_{i+1} \geq u$. For the sake of notation, we define $E_{i,u} = \{X_1, \dots, X_i < u\}$ for $i = 0, \dots, n-2$. We need to compute $\mathbb{P}(E_{i,u} \mid E)$ and $\mathbb{P}(X_{i+1} < u' \mid E_{i,u}, E)$ for $u' \geq u$.

Clearly $\mathbb{P}(X_1 < u, \dots, X_i < u \mid E) = 1$ for $i = 0$ so let's assume that $i > 0$. Then,

$$\begin{aligned} \mathbb{P}(E_{i,u} \mid E) &= n \int_0^\infty \mathbb{P}(X_1, \dots, X_i < u, X_1, \dots, X_{n-1} < x) dF(x) \\ &= n \int_0^u F(x)^{n-1} dF(x) + n \int_u^\infty F(u)^i F(x)^{n-i-1} dF(x) \\ &= F(u)^n + nF(u)^i \left(\frac{1 - F(u)^{n-i}}{n-i} \right) \end{aligned}$$

Note that if we assume that $0^0 = 1$, then the formula above also work in the case $i = 0$.

Now, for $u' \geq u$,

$$\mathbb{P}(X_{i+1} < u' \mid E_{i,u}, E) = n \frac{\mathbb{P}(X_{i+1} < u', E_{i,u}, E)}{\mathbb{P}(E_{i,u} \mid E)}$$

we already computed the denominator; hence, we focus on computing the numerator.

$$\begin{aligned} n\mathbb{P}(X_{i+1} < u', E_{i,u}, E) &= n \int_0^\infty \mathbb{P}(X_{i+1} < u', X_1, \dots, X_i < u, X_1, \dots, X_{n-1} < x) dF(x) \\ &= n \int_0^u F(x)^{n-1} dF(x) + n \int_u^{u'} F(u)^i F(x)^{n-i-1} dF(x) \\ &\quad + nF(u)^i F(u') \int_{u'}^\infty F(x)^{n-i-2} dF(x) \\ &= F(u)^n + nF(u)^i \left(\frac{F(u')^{n-i} - F(u)^{n-i}}{n-i} \right) \\ &\quad + nF(u)^i F(u') \left(\frac{1 - F(u')^{n-i-1}}{n-i-1} \right) \end{aligned}$$

$$\begin{aligned}
&= F(u)^n - \frac{n}{n-i} F(u)^n + \frac{n}{n-i} F(u)^i F(u')^{n-i} \\
&\quad + \frac{n}{n-i-1} F(u)^i F(u') - \frac{n}{n-i-1} F(u)^i F(u')^{n-i}
\end{aligned}$$

Then,

$$\frac{d}{dx} \mathbb{P}(X_{i+1} < x \mid E_{i,u}, E) = F(u)^i \left(\frac{n}{n-i-1} \right) \frac{1 - F(x)^{n-i-1}}{\mathbb{P}(E_{i,u} \mid E)} \frac{dF}{dx}(x)$$

and

$$\mathbf{E} [X_{i+1} \mathbf{1}_{\{X_{i+1} \geq u\}} \mid E_{i,u}, E] = \frac{1}{\mathbb{P}(E_{i,u} \mid E)} F(u)^i \frac{n}{n-i-1} \int_u^\infty x (1 - F(x)^{n-i-1}) dF(x).$$

Then,

$$\begin{aligned}
\mathbb{E}(ALG) &= \sum_{i=0}^{n-2} \frac{n}{n-i-1} F(u)^i \int_u^\infty x (1 - F(x)^{n-i-1}) dF(x) \\
&= \sum_{i=0}^{n-2} \frac{n}{n-i-1} (1-q)^i \int_0^q F^{-1}(1-w)(1 - (1-w)^{n-i-1}) dw \\
&\quad \text{(Change of variable } 1-q = F(u)) \\
&= \sum_{i=0}^{n-2} \frac{n}{n-i-1} (1-q)^i \int_0^q \int_w^1 r(v) dv (1 - (1-w)^{n-i-1}) dw \\
&\quad \text{(Using that } F^{-1}(1-w) \text{ is strictly decreasing and differentiable)} \\
&= \int_0^1 r(v) \sum_{i=0}^{n-2} \frac{n}{n-i-1} (1-q)^i \left(\min\{q, v\} - \left(\frac{1 - (1 - \min\{q, v\})^{n-i}}{n-i} \right) \right) dv
\end{aligned}$$

C.2 Proof of Proposition 3

Proof of Proposition 3. For $v > q$, we have

$$A_{n,q}(v) = \frac{\sum_{i=0}^{n-2} \frac{n}{n-i-1} (1-q)^i \left(q - \left(\frac{1 - (1-q)^{n-i}}{n-i} \right) \right)}{\mathbb{P}(\text{Binom}(n, v) \geq 2)} = \frac{1 - (1-q)^{n-1} (1 + (n-1)q)}{\mathbb{P}(\text{Binom}(n, v) \geq 2)}$$

This last function is decreasing in v attaining its minimum at $v = 1$.

For $v \leq q$, we have

$$\begin{aligned}
A_{n,q}(v) &= \frac{\sum_{i=0}^{n-2} \frac{n}{n-i-1} (1-q)^i \left(v - \left(\frac{1 - (1-v)^{n-i}}{n-i} \right) \right)}{\mathbb{P}(\text{Binom}(n, v) \geq 2)} \\
&= \sum_{i=0}^{n-2} \frac{n}{(n-i-1)(n-i)} (1-q)^i \left(\frac{(1-v)^{n-i} - (1 - (n-i)v)}{\mathbb{P}(\text{Binom}(n, v) \geq 2)} \right)
\end{aligned}$$

$$= \sum_{i=2}^n \frac{n}{i(i-1)} (1-q)^i G_{n,n-i}(v),$$

where $G_{n,i}(v) = \frac{(1-v)^i - (1-iv)}{\mathbb{P}(\text{Binom}(n,v) \geq 2)}$ for $i \in \{2, \dots, n\}$. To conclude that $A_{n,q}(v)$ is increasing in v , it is enough to show that $G_{n,i}(v)$ is increasing in v for all $i \in \{2, \dots, n\}$. Then,

$$G'_{n,i}(v) = \frac{g_{n,i}(v)}{\mathbb{P}(\text{Binom}(n,v) \geq 2)^2},$$

where $g_{n,i}(v) = (-i(1-v)^{i-1} + i)\mathbb{P}(\text{Binom}(n,v) \geq 2) - ((1-v)^i - (1-iv))n(n-1)v(1-v)^{n-2}$. To conclude the proof, it is enough to show that $g_{n,i}(v) \geq 0$. We note that $g_{n,i}(0) = 0$, so we only need to prove that $g'_{n,i}(v) \geq 0$. Now,

$$\begin{aligned} g'_{n,i}(v) &= i(i-1)(1-v)^{i-2}\mathbb{P}(\text{Binom}(n,v) \geq 2) - ((1-v)^i - (1-iv))n(n-1)(1-v)^{n-2} \\ &\quad + ((1-v)^i - (1-iv))n(n-1)(n-2)v(1-v)^{n-3} \\ &= i(i-1)(1-v)^{i-2}\mathbb{P}(\text{Binom}(n,v) \geq 2) \\ &\quad - n(n-1)(1-(n-1)v)(1-v)^{n-3}((1-v)^i - (1-iv)) \end{aligned}$$

Using the second equality it is easy to verify that $g'_{n,i}(v) \geq 0$ for $v \geq 1/(n-1)$. Hence, from now, we assume that $v < 1/(n-1)$. Furthermore, by inspection, we can verify that $g'_{n,i}(v) \geq 0$ for $n \in \{3, 4\}$ and $i \in \{2, \dots, n\}$; hence, from now on, we assume that $n \geq 5$. The following claim allows us to focus only on lower bounding $g'_{n,2}(v)$.

Claim 1. *For $n \geq 5$, $v \leq 1/(n-1)$ and for all $i \in \{2, \dots, n-1\}$, we have $g'_{n,i}(v) \leq g'_{n,i+1}(v)$.*

This proof requires lower bounding several polynomials and various case analysis; hence, we defer it to the end. Now, note that

$$\begin{aligned} g'_{n,2}(v) &= 2\mathbb{P}(\text{Binom}(n,v) \geq 2) - n(n-1)(1-(n-1)v)(1-v)^{n-3}v^2 \\ &\geq n(n-1)v^2(1-v)^{n-2} - n(n-1)(1-(n-1)v)v^2(1-v)^{n-3} \\ &= n(n-1)v^2(1-v)^{n-3}(1-(1-v)(1-(n-1)v)) \\ &= n(n-1)v^2(1-v)^{n-3}(v + (n-1)v(1-v)) \end{aligned}$$

where in the first inequality we use the lower bound $\mathbb{P}(\text{Binom}(n,v) \geq 2) \geq \mathbb{P}(\text{Binom}(n,v) = 2)$. From here, we obtain that $g'_{n,2}(v) \geq 0$ with equality at $v = 0$. Using Claim 1, we conclude that $G'_{n,i}(v) \geq 0$ for all $i \in \{2, \dots, n\}$.

Proof of Claim 1. Indeed,

$$\begin{aligned} g'_{n,i}(v) - g'_{n,i+1}(v) &= -\mathbb{P}(\text{Binom}(n,v) \geq 2) (2 - (i+1)v) i (1-v)^{i-2} \\ &\quad + v(1 - (1-v)^i)(1 - (n-1)v)n(n-1)(1-v)^{n-3} \\ &\leq (1-v)^{i-3} \left(-i(2 - (i+1)v)(1-v)\mathbb{P}(\text{Binom}(n,v) \geq 2) \right. \end{aligned} \tag{C6}$$

$$+ v(1 - (1 - v)^i)(1 - (n - 1)v)n(n - 1)) \quad (\text{C7})$$

where in the inequality we use that $(1 - v)^{n-3} \leq (1 - v)^{i-3}$. Now, let $\phi_{n,i}(v)$ be the term in the big parenthesis in (C7), so $g'_{n,i}(v) - g'_{n,i+1}(v) \leq (1 - v)^{i-3}\phi_{n,i}(v)$. We now focus on proving that $\phi_{n,i}(v) \leq 0$. For this, we will reduce the problem to bounding only $\phi_{n,2}(v) \leq 0$ by proving that $\phi_{n,2}(v) \geq \phi_{n,i}(v)$ for all $i = 2, \dots, n - 1$. To prove this last inequality, we analyze the difference $\phi_{n,i}(v) - \phi_{n,i+1}(v)$ for $i \in \{2, \dots, n - 2\}$:

$$\begin{aligned} & \phi_{n,i}(v) - \phi_{n,i+1}(v) \\ &= 2(1 - v)(1 - (i + 1)v)\mathbb{P}(\text{Binom}(n, v) \geq 2) - v^2(1 - v)^i(1 - (n - 1)v)n(n - 1) \\ &\geq v^2(1 - v)^i(n - 1)n \left((1 - (i + 1)v)(1 - v)^{n-i-2} \left(1 + \frac{n-5}{3}v \right) - (1 - v(n - 1)) \right) \\ &= v^2(1 - v)^i(n - 1)n \cdot \theta_{n,i}(v). \end{aligned}$$

Note that the function $(1 - (i + 1)v)(1 - v)^{n-i-2}$ is decreasing in i ; hence, for $n \geq 5$, we have $\theta_{n,i}(v) \geq \theta_{n,n-2}(v)$. From here, we obtain

$$\phi_{n,i}(v) - \phi_{n,i+1}(v) \geq v^2(1 - v)^i(n - 1)n \cdot \theta_{n,n-2}(v).$$

On the other hand, we have $\theta_{n,n-2}(v) = (1 - (n - 1)v) \left(\frac{n-5}{3}v \right) \geq 0$. From here, we obtain that $\phi_{n,i}(v) \geq \phi_{n,i+1}(v)$ for all $i \in \{2, \dots, n - 2\}$ and so $\phi_{n,2}(v) \geq \phi_{n,i}(v)$ for all $i \in \{2, \dots, n - 2\}$. Now,

$$\begin{aligned} \phi_{n,2}(v) &= -2(2 - 3v)(1 - v)\mathbb{P}(\text{Binom}(n, v) \geq 2) + v^2(2 - v)(1 - (n - 1)v)n(n - 1) \\ &\leq -2(2 - 3v)(1 - v) \left(\binom{n}{2}v^2(1 - v)^{n-2} + \binom{n}{3}v^3(1 - v)^{n-3} \right) \\ &\quad + v^2(2 - v)(1 - (n - 1)v)n(n - 1) \\ &= n(n - 1)v^2 \left(-(2 - 3v) \left((1 - v)^{n-1} + \frac{n-2}{3}v(1 - v)^{n-2} \right) + (2 - v)(1 - (n - 1)v) \right) \\ &= n(n - 1)v^2 \left((2 - v)(1 - (n - 1)v) - (2 - 3v)(1 - v)^{n-2} \left(1 + \frac{n-5}{3}v \right) \right) \end{aligned}$$

We analyze this last bound for the case $n = 5$ and case $n \geq 6$ separately. For $n = 5$, we have

$$\phi_{5,2}(v) \leq 20v^2 \left((2 - v)(1 - 4v) - (2 - 3v)(1 - v)^3 \right).$$

The polynomial $(2 - v)(1 - 4v) - (2 - 3v)(1 - v)^3$ has roots $v \in \{0, (11 - i\sqrt{11})/6, (11 + i\sqrt{11})/6\}$ with 0 having multiplicity 2. Since, the polynomial tends to $-\infty$ when $v \rightarrow \infty$ and its only 0 when $v = 0$, we deduce that $\phi_{5,2}(v) \leq 0$.

Now, assume that $n \geq 6$, then

$$\phi_{n,2}(v) \leq n(n - 1)v^2 \left((2 - v)(1 - v)^{n-1} - (2 - 3v)(1 - v)^{n-2} \left(1 + \frac{n-5}{3}v \right) \right)$$

$$= n(n-1)v^2 \left((1-v)^{n-2} \cdot v \cdot \left(-\frac{2(n-5)}{3} + (n-4)v \right) \right)$$

where in the first inequality we use Bernoulli's inequality on $1 - (n-1)v \leq (1-v)^{n-1}$ and in the equality we simply reorder the big parenthesis from the previous line. Now, the polynomial $v(-2(n-5)/3 + (n-4)v)$ has 2 roots at $v \in \{0, 2(n-5)/(3(n-4))\}$. Hence, for $v \leq 1/(n-1)$, we must have that $\phi_{n,2}(v) \leq 0$.

Going back to the function $g'_{n,i}$, all our calculations give us

$$g'_{n,i}(v) - g'_{n,i+1}(v) \leq (1-v)^{i-3} \phi_{n,2}(v) \leq 0,$$

which finishes the proof of Claim 1. □

□

C.3 An Upper Bound for Single-Threshold Solutions

We present an instance that shows that no strategy in the class of single-threshold (including randomization) can obtain a competitive ratio larger than 0.5463. We use a counterexample motivated by [41]. For $n \geq 1$, we consider the following function from $(0, 1]$ to \mathbf{R}_+

$$f(u) = \frac{a \cdot c_n}{u} \mathbf{1}_{(0, 1/n^{10})}(u) + b \cdot \mathbf{1}_{[1/n^{10}, \beta/n]}(u)$$

where $\mathbf{1}_X$ is the indicator function that is 1 for $u \in X$ and 0 for $u \notin X$, $a, b > 0$ and $\beta > 1/n$ are positive constants to be optimized and $c_n = \left(n \cdot \left(1 - (1 - 1/n^{10})^{n-1} \right) \right)^{-1}$. We are going to assume that $a + b \leq 1$. For n large enough, we have that f is nonincreasing.

Now, we can construct a random variable from f as follows. First, we add a small perturbation to f so f is smooth and strictly decreasing. This can be done by taking a convolution with a smooth function. Let's call f_ε the resulting function, with small error $\varepsilon > 0$; hence, when $\varepsilon \rightarrow 0$, we have $f_\varepsilon(u) \rightarrow f(u)$, except for a set of measure 0. Note that f_ε is surjective in \mathbf{R}_+ . Now, for $x \geq 0$, let $F_\varepsilon(x) = 1 - f_\varepsilon^-(x)$. Note that F is increasing, $F_\varepsilon(0) = 0$ and $F_\varepsilon(+\infty) = 1$; hence, F_ε is a valid CDF. We define the random variable X_ε to be the random variable following F_ε . By construction, $F_\varepsilon^{-1}(1-u) = f_\varepsilon(u)$. For $\varepsilon \rightarrow 0$, we have $F_\varepsilon^{-1}(1-u) \rightarrow f(u)$, except for a set of measure 0 in $[0, 1]$. To avoid notational clutter, from now on, we simply work with $f(u)$ instead of f_ε . By abusing notation, we will write $F^{-1}(1-u) = f(u)$, but it has to be understood that this equality occurs except for the points $1/n^{10}$ and β/n .

We now consider a sequence of n independent random variables following F (the limit of F_ε when $\varepsilon \rightarrow 0$). We assume that n is large. The result now follows from the following two Lemmas.

Lemma 11. *We have $\mathbb{E}(X_{(2)}) \rightarrow a + b(1 - e^{-\beta}(1 + \beta))$ when $n \rightarrow \infty$.*

Lemma 12. *There is $n_0 \geq 0$ such that for any algorithm ALG, if the input is of length $n \geq n_0$, the value collected by the algorithm is bounded as*

$$\mathbb{E}(\text{ALG}) \leq p(a, b, \beta) + 5 \frac{\beta(1 + \beta)^2}{n - \beta},$$

where $p(a, b, \beta) = \max_{\lambda \in [0, \beta]} \{a(1 - e^{-\lambda})/\lambda + b(1 - e^{-\lambda}(1 + \lambda))\}$.

We first provide the tight upper bound and then we prove the lemmas. Using these two lemma, for any algorithm and $n \geq n_0$, we have

$$\frac{\mathbb{E}(\text{ALG})}{\mathbb{E}(X_{(2)})} \leq \frac{p(a, b, \beta) + 5\beta(1 + \beta)^2/(n - \beta)}{\mathbb{E}(X_{(2)})}$$

Using numerical optimization to minimize $p(a, b, \beta)$, we found $a \approx 0.5463$, $b \approx 0.4537$ and $\beta \approx 109.131$, we obtain $p(a, b, \beta) \approx 0.5463$. Hence, for n large, we obtain that $\mathbb{E}(\text{ALG})/\mathbb{E}(X_{(2)}) \leq 0.5463 + o(n)$. This shows that with one threshold, we cannot recover the approximation of $1 - 1/e \approx 0.6321$ in the standard prophet inequality.

In the remainder of the subsection, we present the proof of Lemma 11 and 12.

Proof of Lemma 11. We have

$$\begin{aligned} \mathbb{E}(X_{(2)}) &= n(n-1) \int_0^1 F^{-1}(1-u)q(1-q)^{n-2} dq \\ &= n(n-1) \int_0^{1/n^{10}} a \cdot c_n(1-q)^{n-2} dq + n(n-1) \int_{1/n^{10}}^{\beta/n} bq(1-q)^{n-2} dq \\ &= ac_n n \left(1 - \left(1 - \frac{1}{n^{10}}\right)\right) \\ &\quad + b \left(\frac{1}{n^9} \left(1 - \frac{1}{n^{10}}\right)^{n-1} - \beta \left(1 - \frac{\beta}{n}\right)^{n-1} + \left(1 - \frac{1}{n^{10}}\right)^n - \left(1 - \frac{\beta}{n}\right)^n\right) \end{aligned}$$

The conclusion now follows by taking limit in the last equality. \square

Proof of Lemma 12. We can parametrize every single-threshold algorithm via the quantile chosen by it. If ALG_q denotes the value obtained by a single-threshold algorithm that always uses quantile q , we have $\text{ALG} \leq \max_{q \in [0, 1]} \text{ALG}_q$. We analyze this last maximum for $q \leq 1/n^{10}$, $q \in [1/n^{10}, \beta/n]$ and $q \geq \beta/n$.

For $q \leq 1/n^{10}$, we have

$$\begin{aligned} \mathbb{E}(\text{ALG}_q) &= \sum_{k=1}^{n-1} \frac{n}{k} (1-q)^{n-k-1} \int_0^q \frac{ac_n}{w} (1 - (1-w)^k) dw \\ &\leq a \cdot c_n \sum_{k=1}^{n-1} \frac{n}{k} \sum_{\ell=0}^{k-1} \int_0^{1/n^{10}} (1-w)^\ell dw \end{aligned}$$

$$\begin{aligned}
&= ac_n \sum_{k=1}^{n-1} \sum_{\ell=0}^{k-1} \frac{1 - (1 - 1/n^{10})^{\ell+1}}{\ell + 1} \\
&\leq ac_n \sum_{k=1}^{n-1} \frac{n}{k} \frac{k}{n^{10}} \\
&= a \cdot \frac{1/n^9}{1 - (1 - 1/n^{10})^{n-1}} \leq a \left(1 + \frac{3}{n-1}\right)
\end{aligned}$$

where in the first inequality we upper bounded the integral for $q = 1/n^{10}$ and we also upper bounded $(1-q)^{n-k-1} \leq 1$; in the second equality we performed the integration; in the second inequality we used Bernoulli's inequality: $(1-1/n^{10})^{\ell+1} \geq 1 - (\ell+1)/n^{10}$, and in the last inequality we used the following claim.

Claim 2. *We have $c_n/n^8 \leq (1 + 3/(n-1))$.*

Proof. First, note that

$$\left(1 - \frac{1}{n^{10}}\right)^{n-1} \leq e^{-(n-1)/n^{10}} \leq 1 - \frac{n-1}{n^{10}} + \frac{(n-1)^2}{2n^{20}} \quad (\text{C8})$$

Then,

$$\begin{aligned}
\frac{c_n}{n^8} &= \frac{1/n^9}{1 - (1 - 1/n^{10})^{n-1}} \\
&\leq \frac{n}{(n-1)(1 - (n-1)/(2n^{10}))} \\
&\leq \left(1 + \frac{1}{n-1}\right) \left(1 + \frac{1}{2n}\right) \\
&\leq 1 + \frac{3}{n-1}
\end{aligned}$$

where in the first inequality we used inequality (C8), the second inequality follows by $1 \leq (1 - (n-1)/2n^{10})(1 + 1/2n)$ and the last inequality follows simply by expanding the multiplication and bounding $1/2n, 1/(2n(n-1)) \leq 1/(n-1)$. \square

Therefore, for any $q \leq 1/n^{10}$, we have that the value of the algorithm is upper bounded by $a \cdot (1 + 3/(n-1))$.

For $q = \lambda/n \in [1/n^{10}, \beta]$, we have

$$\begin{aligned}
\mathbb{E}(\text{ALG}_q) &= \sum_{k=1}^{n-1} \frac{n}{k} (1-q)^{n-k-1} \int_0^{1/n^{10}} \frac{ac_n}{w} (1 - (1-w)^k) dw \\
&\quad + \sum_{k=1}^{n-1} \frac{n}{k} (1-q)^{n-k-1} \int_{1/n^{10}}^q b (1 - (1-w)^k) dw
\end{aligned}$$

$$\begin{aligned}
&\leq ac_n \sum_{k=1}^{n-1} \frac{n}{k} (1-q)^{n-k-1} \frac{k}{n^{10}} \\
&\quad + b \sum_{k=1}^{n-1} \frac{n}{k} (1-q)^{n-k-1} \left(\frac{(1-q)^{k+1} - ((1-1/n^{10})^{k+1} - (q-1/n^{10})(k+1))}{k+1} \right)
\end{aligned}$$

where in the first inequality we upper bounded the first term in the first line in a manner similar to the case $q \leq 1/n^{10}$ and we integrated the second term. The following two claims allow us to upper bound this last inequality by controlling the error.

Claim 3. We have $c_n (1 - (1-q)^{n-1}) / (qn^9) \leq (1 - e^{-\lambda}) / \lambda + \beta(1 + \beta) / (n - \beta)$

Proof. In the proof, we use that $\lambda \leq \beta \leq 2$. First, we note that

$$\begin{aligned}
1 - \left(1 - \frac{\lambda}{n}\right)^{n-1} &= 1 - e^{-\lambda} + e^{-\lambda} - \left(1 - \frac{\lambda}{n}\right)^{n-1} \\
&\leq 1 - e^{-\lambda} + e^{-\lambda} - e^{-\lambda \left(\frac{n-1}{n-\lambda}\right)} \\
&= 1 - e^{-\lambda} + e^{-\lambda} \left(1 - e^{\lambda \left(\frac{1-\lambda}{n-\lambda}\right)}\right) \\
&\leq 1 - e^{-\lambda} + \frac{|\lambda(1-\lambda)|}{n-\lambda} \leq 1 - e^{-\lambda} + \frac{\beta(1+\beta)}{n-\beta}
\end{aligned}$$

where in the first inequality we used that $1/(1-\lambda/n) \leq e^{\lambda/(n-\lambda)}$ using the standard inequality $1+x \leq e^x$, in the second inequality we used that $e^{-\lambda} \leq 1$ and $1-|x| \leq e^{-x}$, and in the last inequality, we simply used that $\lambda \leq \beta$. \square

Claim 4. We have

$$\begin{aligned}
&\sum_{k=1}^{n-1} \frac{n}{k} (1-q)^{n-k-1} \left(\frac{(1-q)^{k+1} - ((1-1/n^{10})^{k+1} - (q-1/n^{10})(k+1))}{k+1} \right) \\
&\leq 1 - e^{-\lambda}(1+\lambda) + 3 \frac{\beta(1+\beta)^2}{n}.
\end{aligned}$$

Proof. First, we have

$$\begin{aligned}
\sum_{k=1}^{n-1} \frac{n}{k(k+1)} (1-q)^{n-k-1} ((1-q)^{k+1} - (1-q(k+1))) &= \sum_{k=1}^{n-1} nq^2 (1-q)^{n-k-1} \\
&\quad - (1-q)^n + (1-qn) \\
&= 1 - (1-q)^{n-1} (1-q+qn)
\end{aligned}$$

where in the first equality we simply rearranged the sum and in the next line we computed the sum. Then,

$$\begin{aligned}
& \sum_{k=1}^{n-1} \frac{n}{k} (1-q)^{n-k-1} \left(\frac{(1-q)^{k+1} - ((1-1/n^{10})^{k+1} - (q-1/n^{10})(k+1))}{k+1} \right) \\
& \leq \sum_{k=1}^{n-1} \frac{n}{k(k+1)} (1-q)^{n-k-1} ((1-q)^{k+1} - (1-q(k+1))) \\
& \quad + \sum_{k=1}^{n-1} \frac{n}{k(k+1)} (1-q)^{n-k-1} (1 - (1-1/n^{10})^{k+1}) \\
& \leq 1 - (1-q)^{n-1} (1-q+qn) + \sum_{k=1}^{n-1} \frac{1}{k} \frac{1}{n^9} \\
& \leq 1 - (1-q)^{n-1} (1-q+qn) + \frac{\ln(n)}{n^9}
\end{aligned}$$

where in the second inequality, we added and subtracted 1 to the parenthesis to form the term studied at the beginning of the proof, in the second inequality, we used Bernoulli's inequality and in the last inequality, we used the standard Harmonic sum bound. Finally,

$$\begin{aligned}
-(1-q)^{n-1} (1-q+qn) &= -(1+\lambda) \left(1 - \frac{\lambda}{n} \right)^{n-1} + \frac{\lambda}{n} \left(1 - \frac{\lambda}{n} \right)^{n-1} \\
&\leq -(1+\lambda) e^{-\lambda \left(\frac{n-1}{n-\lambda} \right)} + \frac{\beta}{n} \\
&= \left(e^{-\lambda} - e^{-\lambda \left(\frac{n-1}{n-\lambda} \right)} \right) (1+\lambda) - e^{-\lambda} (1+\lambda) + \frac{\beta}{n} \\
&\leq -e^{-\lambda} (1+\lambda) + \frac{\beta(1+\beta)^2}{n-\beta} + \frac{\beta}{n}
\end{aligned}$$

where in the second and last inequalities, we used the same bounds used in the analysis of Claim 3. From here, the inequality of the claim follows. \square

Therefore,

$$\mathbb{E}(\text{ALG}_q) \leq a \cdot \left(\frac{1-e^{-\lambda}}{\lambda} \right) + b(1-e^{-\lambda}(1+\lambda)) + 4 \frac{\beta(1+\beta)^2}{n-\beta}$$

For $q \geq \beta/n$, we have

$$\mathbb{E}(\text{ALG}_q) = \sum_{k=1}^{n-1} \frac{n}{k} (1-q)^{n-k-1} \int_0^{1/n^{10}} \frac{ac_n}{w} (1-(1-w)^k) dw$$

$$+ \sum_{k=1}^{n-1} \frac{n}{k} (1-q)^{n-k-1} \int_{1/n^{10}}^{\beta/n} b(1-(1-w)^k) \, dw.$$

This last term is decreasing in q ; hence it attains its maximum at $q = \beta/n$.

The conclusion of the lemma follows by putting together the three bounds that we found. Additionally, the bound for $\lambda \in [1/n^9, \beta]$ supersedes the bound for $q \leq 1/n^{10}$ and $q \geq \beta/n$. \square